DOCUMENT RESUME

ED 279 725 TM 870 174

AUTHOR Fingerman, Paul W.
TITLE Recommendations for Editing the Common Core of Data, Parts VI and VI-A. Technical Report No. 11.
INSTITUTION American Institutes for Research in the Behavioral Sciences. Palo Alto, CA. Statistical Analysis Group in Education.
SPONS AGENCY National Center for Education Statistics (DHEW), Washington, D.C.
PUB DATE Jul 79
CONTRACT 300-78-0150
NOTE 88p.; Some pages are marginally legible.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Data Analysis; Data Collection; *Data Processing; Elementary Secondary Education; *National Surveys; Public Schools; School Districts; *School Surveys; Statistical Analysis; Statistical Data; *Statistical Surveys
IDENTIFIERS National Center for Education Statistics

ABSTRACT

Recommendations are made for data processing procedures for the National Center for Education Statistics' (NCES) Common Core of Data survey program. Data are anticipated from about 16,000 local education agencies and 85,000 schools in the 50 states, the District of Columbia, and 7 territories. These recommendations concern editing and verification performed between the receipt of data and the start of substantive analyses--inspecting for suspect values, as well as detecting and correcting erroneous values. A flow chart details the editing process, which includes online data entry edits, preliminary machine edits, and batch production edits. The precise nature and relationship between the data entry and preliminary edit phase depends on the form in which the data are received. Online data entry edits are strongly recommended, since the first check will usually be for a keying error and the source document is still immediately available. The preliminary edit examines the data statistically, searching for frequent or universal format-type problems, while the production edit attempts to individually identify every field or record that fails an edit. Other described processes include automatic correction, relational or longitudinal edits, table-driven edits, switch-driven edits, and user-oriented interfaces. Specific suggestions are described for two surveys: the Universe of Public Schools Survey and the Local Education Agency Non-Fiscal Survey. Various survey forms and data processing instructions are appended. (GDC)

Technical Report No. 11

Recommendations for Editing the Common Core of Data Parts VI and VI-A

Paul W. Fingerman

Prepared by

SAGE
STATISTICAL ANALYSIS GROUP IN EDUCATION

For the

National Center for Education Statistics

American Institutes for Research

Box 1113, Palo Alto, California 94302

TECHNICAL REPORT 11

RECOMMENDATIONS FOR EDITING THE COMMON CORE OF DATA, PARTS VI AND VI-A

Paul W. Fingerman

Statistical Analysis Group in Education
American Institutes for Research
P.O. Box 1113
Palo Alto, California 94302

This work was done under Contract No. 300-78-0150 with the National Center for Education Statistics, Department of Health, Education, and Welfare. However, the content does not necessarily reflect the position or policy of either agency, and no official endorsement should be inferred.

July 1979

PREFACE

This report has been prepared as part of Task 2D of Project SAGE. The edit descriptions in the report will be used as draft specifications for the 1978-79 cycle of the Common Core of Data (CCD) program. Major data-handling systems are constantly undergoing revision, and two factors interacted to stimulate enhancement of the existing system at this time. First, the survey instruments for Parts VI-A and VI of CCD have been frozen until 1981; thus it is possible to concentrate on improving the system, rather than racing this year merely to adapt the programs to new survey items. Second, resources were available, through Project SAGE, to obtain an independent review of the existing system from analysts with their own practical experience in the survey data-processing field. The resulting combination of NCES and SAGE insight will lead to an improved product. Finally, it is appropriate to acknowledge two individuals who were primarily responsible for the edit procedures employed last year, who participated in the analyses which preceded this report, and who provided the foundation for the material developed herein: Warren Hughes and Ted Chmura.

TABLE OF CONTENTS

LIST OF TABLES & FIGURES

TABLES

FIGURES

INTRODUCTION

Parts VI-A and VI of the Common Core of Data survey program are being put through their first full-scale run this year. NCES anticipates receipt of data this year on approximately 16,000 local education agencies (LEAs) and perhaps 85,000 schools from 58 states and territories. While these data promise to be invaluable in assessing the status of education today and in establishing trends for the future, the millions of characters of input that are involved will require attentive processing from the very outset. The purpose of this report is to present recommendations for that processing which is to be performed between the receipt of the data and the start of substantive analyses; this period of data examination, of inspection for suspect values, of detection and correction of erroneous values, is referred to as the "edit and verification" stage of processing.

The next section presents some background and general recommendations regarding the overall structure of the edit system. No attempt is made to review every possible practice in the design of edit systems; rather, only those instances are discussed where the CCD edit system might deviate from more conventional approaches or might benefit from special enhancement. The two remaining sections of the report deal with specific draft edit specifications for Parts VI-A and VI respectively.

SYSTEM-LEVEL CONSIDERATIONS AND RECOMMENDATIONS

Background

According to the system plan for Parts VI-A and VI prepared by SAGE (Figures 1a-1f), edit processing is to include three major components: edits during data entry, preliminary edits, and batch production edits. The precise nature and relationship between the data entry and preliminary edit phase depends on the form in which the data are received (c.f. Figure 1c). At least three media are expected, and preliminary edits are recommended for all. For hard-copy (form or facsimile) and shuttle-list data, both manual and machine preliminary edits must be prepared. A manual/clerical scan is used to separate out forms that are unreadable, are not filled out properly, are missing almost all data items, and so on, and to count the number of input records (i.e., schools or LEAs). Gross problems or problems that would impact on data entry are discovered and corrected at this stage. Data entry (Figure 1c) follows the clerical screening. Inputting the data via CRT terminal, using any of a number of commercially available data entry software packages or services, a number of edit checks should be performed as the data are entered. These include field content verification (e.g., only numbers in numeric fields), presence checks (e.g., non-optional fields, like LEA identification numbers, are correctly filled out), range checks for numerical items, validity checks for coded items (e.g., table look-up to verify codes are legal), internal consistency checks and others. The advantage of doing such checks at the time of data entry is that the source document is immediately available; follow-up is thus facilitated, since the first check will generally be for a keying error. When errors of these types are only discovered later, during a batch edit phase, the source document must be located to check keying, no small task given the volume of input to be processed by CCD.

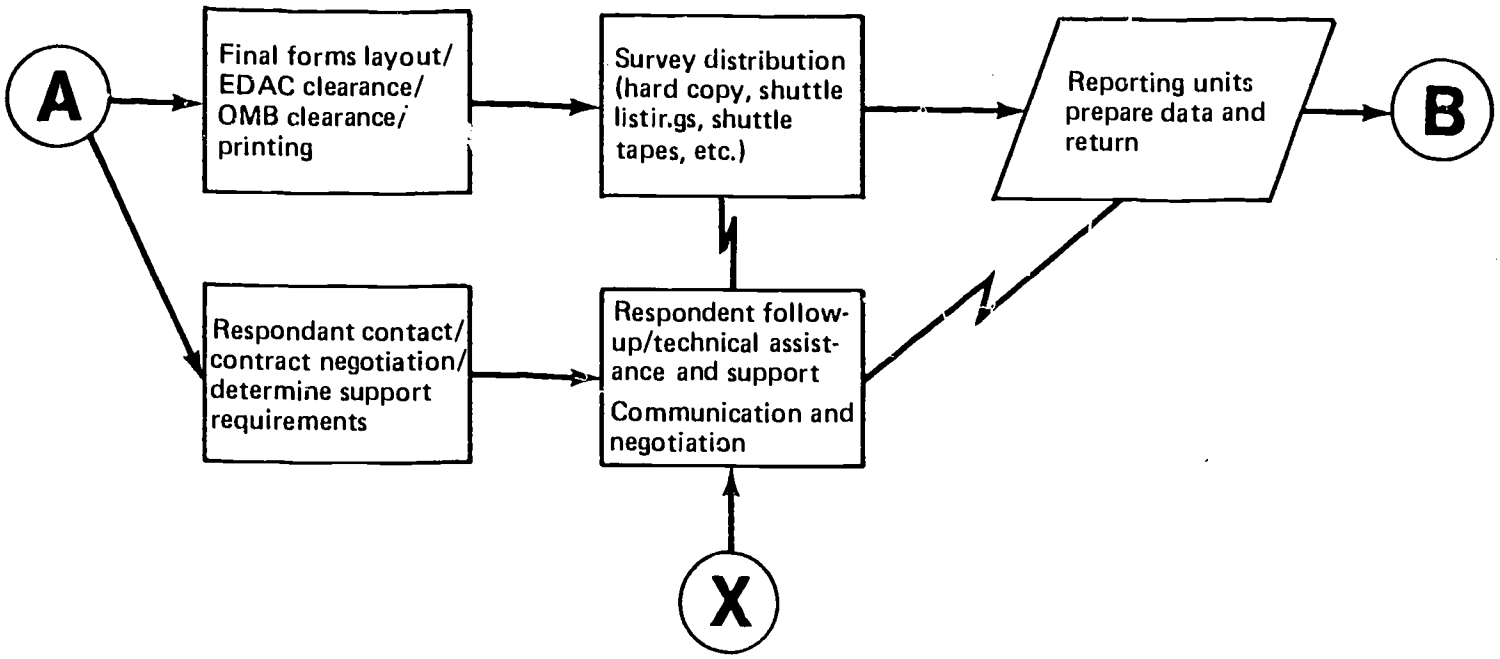

Figure 1a. SYSTEM FLOW CHART - PARTS VI AND VI-A, COMMON CORE OF DATA

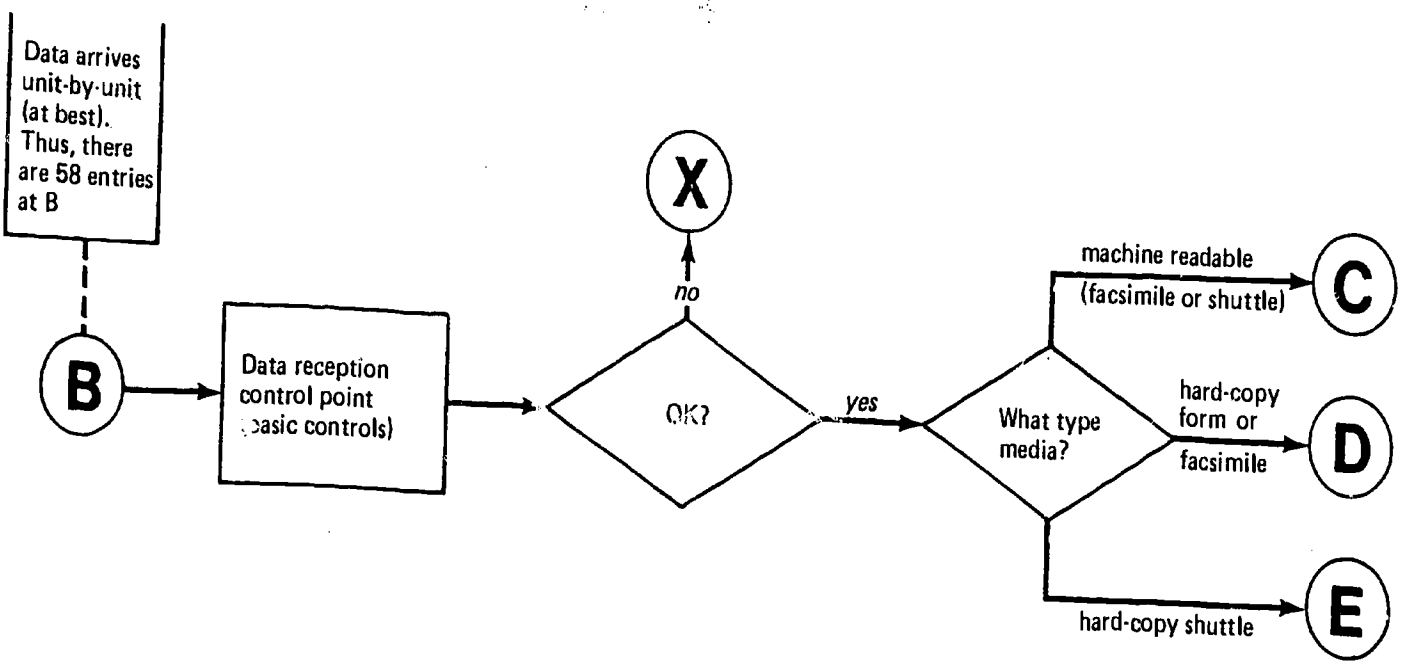

Figure 1b. SYSTEM FLOW CHART - PARTS VI AND VI-A, COMMON CORE OF DATA (cont'd.)

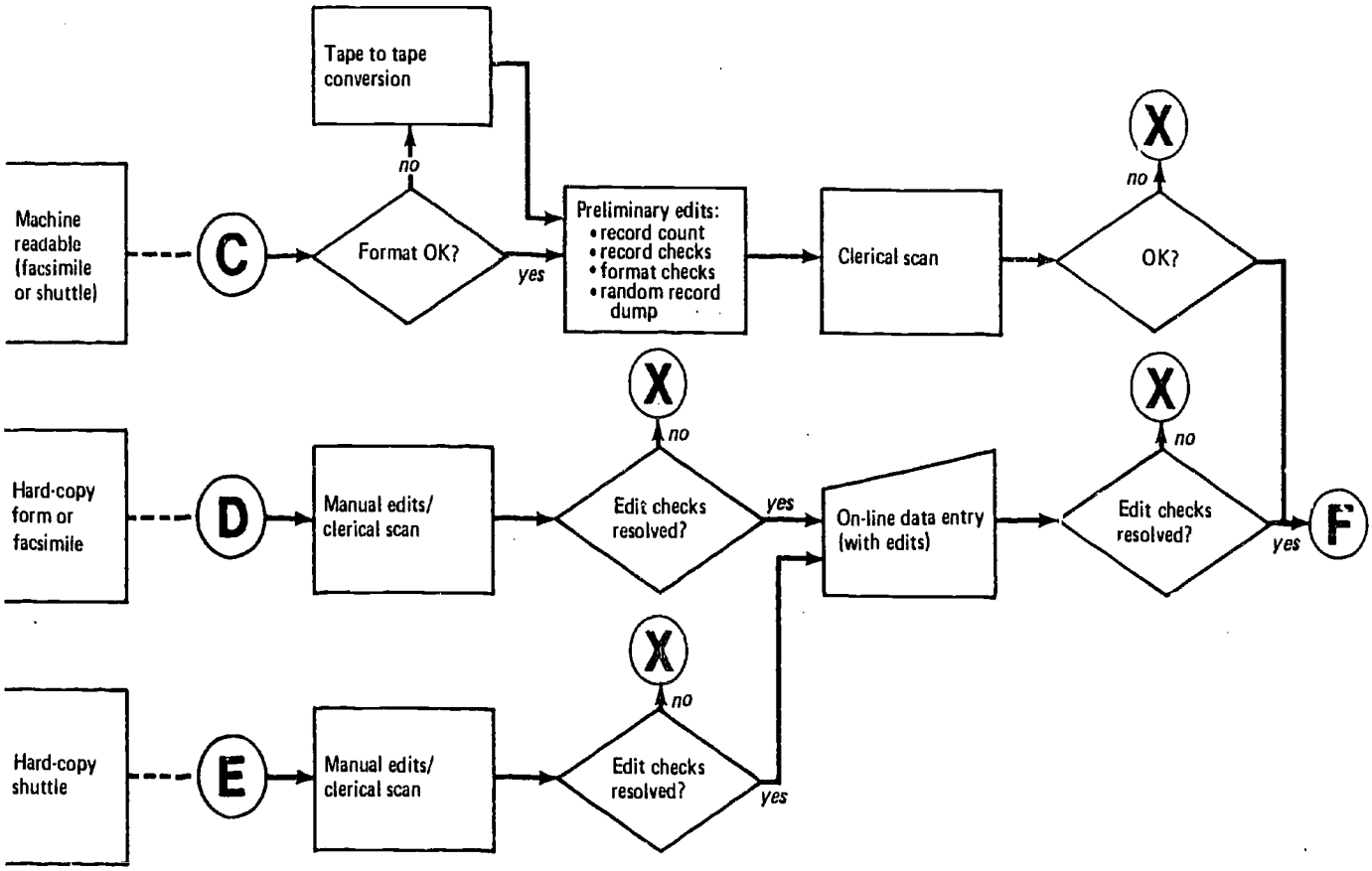

Figure 1c. SYSTEM FLOW CHART - PARTS VI AND VI-A, COMMON CORE OF DATA (cont'd.)

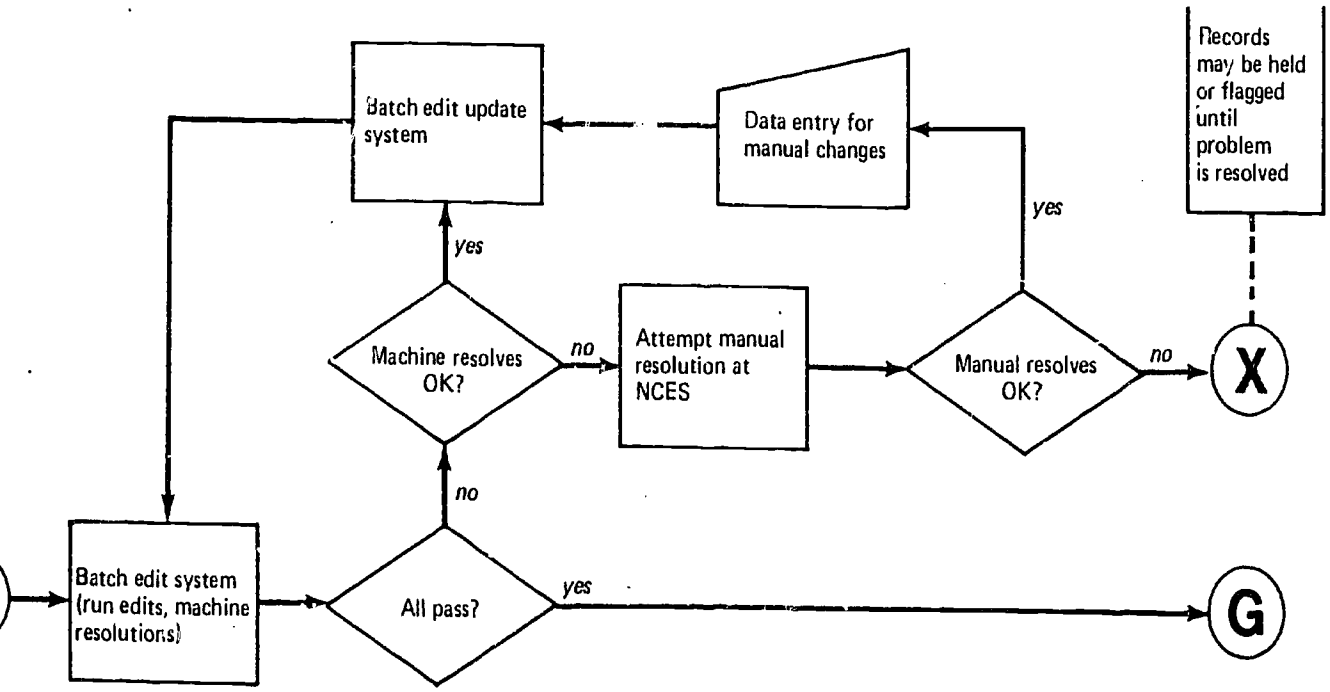

Figure 1d. SYSTEM FLOW CHART - PARTS VI AND VI-A, COMMON CORE OF DATA (cont'd.)

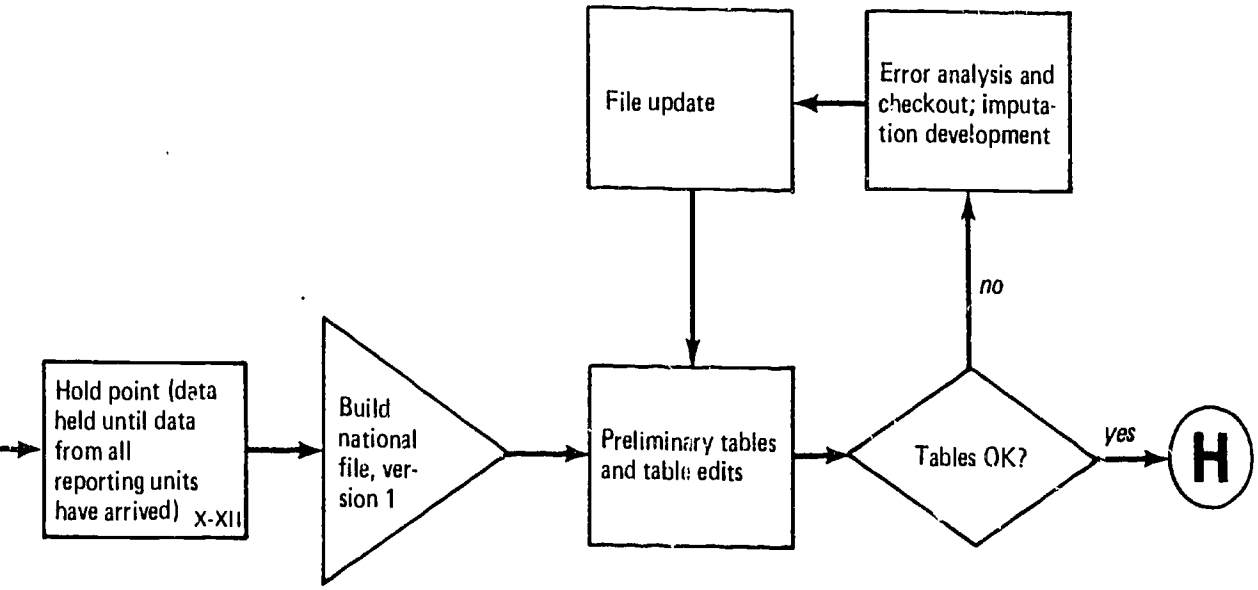

Figure 1e. SYSTEM FLOW CHART - PARTS VI AND VI-A, COMMON CORE OF DATA (cont'd.)

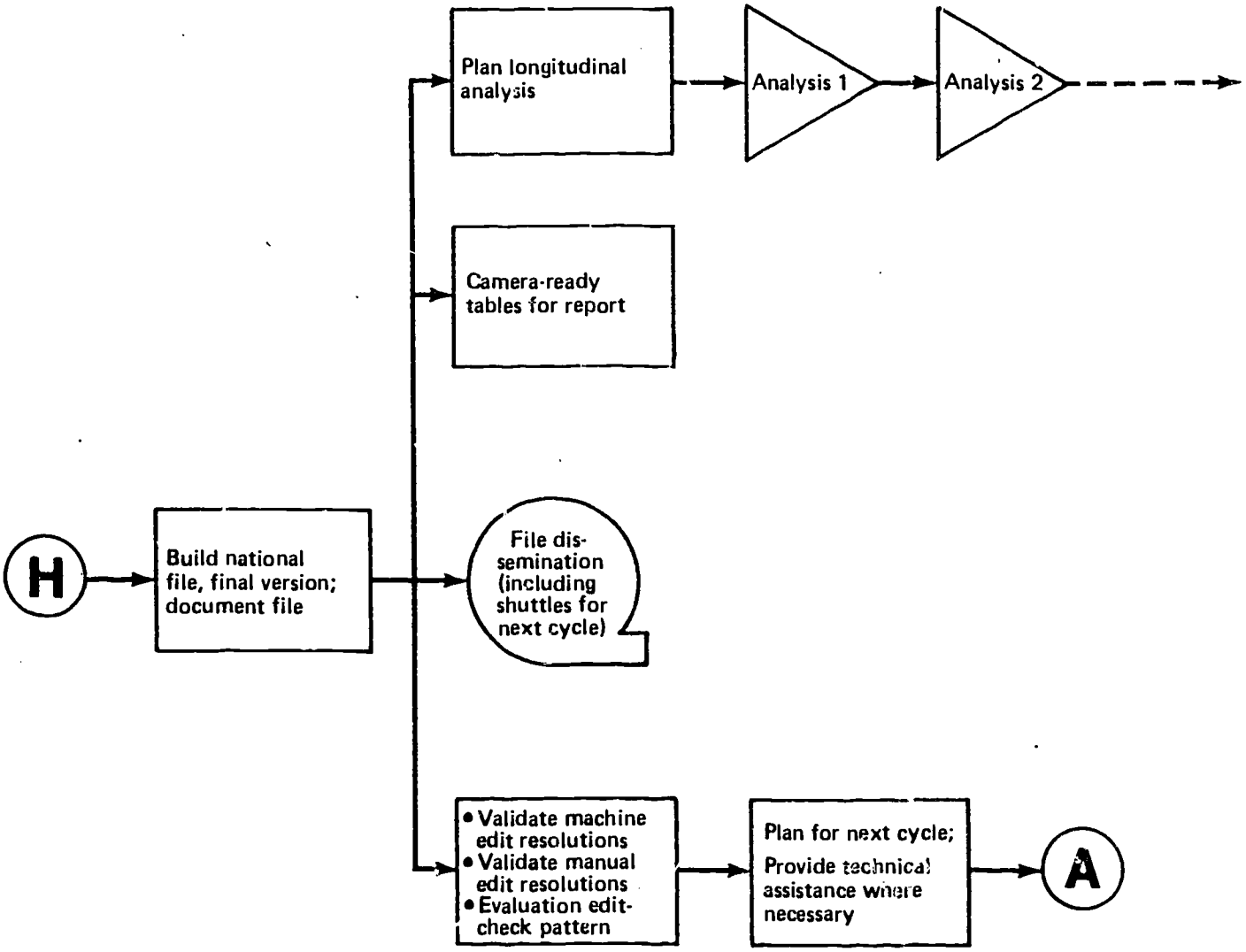

Figure 1f. SYSTEM FLOW CHART - PARTS VI AND VI-A, COMMON CORE OF DATA (cont'd.)

Data entry is followed by the preliminary computer edit designed for machine-readable data.*

A preliminary edit phase is also called for in dealing with data submitted in machine-readable form (Figure 1c). Past experience indicates that several states will likely make systematic errors in preparing data for submission, and that these errors will often cause every (or nearly every) record to be flagged by the production edit program, generating voluminous error printouts. The preliminary machine edit is designed to discover such systematic errors very quickly. For example, in last year's public school universe survey several states used codes for the grade span which were not specified in the instructions (e.g. "K " or " K" for kindergarten, in place of the prescribed "KG"); this led to flagging and lengthy error printouts for hundreds of otherwise correct records. A preliminary edit would have prevented this by providing a frequency count on all entries, before the production edit phase; the use of an alternative code would have been apparent, and a transformation/correction program could have quickly been prepared. Another example is the state that accidentally left blank the field for number of graduates in last year's LEA Non-Fiscal Survey. A preliminary presence check on the field (i.e., a count of the number of LEAs for which it was missing in a given state) would have revealed the problem very quickly.

While the preliminary edit system may check for many of the same problems as are examined in the production edit,

---

*It would be theoretically possible to do all edits proposed in this report at data entry time. This has not been recommended in this report because NCES hopes that most states will ultimately submit their data in machine-readable form. Nevertheless, the possibilities inherent in on-line data entry given current technology are extensive, and should be kept in mind, even as solutions to short-term problems.

there are critical differences: The preliminary edit examines a state's data statistically, searching for frequent or universal, format-type problems; the production edit attempts to individually identify every field and record that fail an edit, and to correct many kinds of edit failures, automatically where possible, and manually when required. Thus, instead of producing a single statistical report for a given state, the production edit program produces an edit report for each record with an error. In some cases this report will indicate an automatic correction has been made (thus permitting a manual override of the correction when necessary); otherwise, sufficient information will be provided to permit a human analyst to provide a correction (via follow-up to the original source of the data at the state level, etc.).

Because of the detect/correct cycle inherent in the production edit phase (Figure 1d), a special file system is also often required. Data files are loaded into the system when first edited, and a temporary expanded version of the file is retained. This intermediate file usually has extra data items called flags to indicate for each data field whether it has passed edit, has failed an edit and awaits correction, has failed an edit and has been automatically corrected, has failed an edit and is to be left unchanged (and ignored by future edit runs on the same file), and so on. This file system is often constructed using a special file access method (e.g., an indexed sequential access method, direct access method, etc.), to facilitate updates to individual records (manually-supplied corrections). Once updated, the intermediate file may be run through the edit process once again (to check updated values, and to check fields which were not editable because other, prerequisite, fields failed the first edit). Each time through the production edit program, the flags are used to determine which fields to edit, which to

ignore, and so on. A final edit run is used to process the file back from its intermediate form (with edit flags) to the final file format.*

Finally, in addition to the special file system, editing requires some method of inserting corrected or updated values. It is often the case that some existing utility or package program (e.g., SAS) is sufficient; occasionally a special purpose program is required for updates in the edit system.

General Recommendations

As indicated in the preface, it was not a major flaw in the existing edit system that led to this report, but rather the availability of resources (time and SAGE) to permit improvement of the system. In this vein, the discussion which follows deals with a number of major and minor edit system components which should be considered in the overall CCD edit process. The remainder of this section deals with proposed improvements in the system, beginning with two major components, data entry and preliminary editing, followed by recommended enhancements to the batch production system.

On-line data entry and edit. This topic has been dealt with above, but a few points are worth pursuing. Many desirable edits can be performed automatically, using commercially available data entry software; the remainder can be programmed and added to most such software packages. The opportunity to conduct edit checks at data entry time is tremendously valuable, since many errors will be found to be due to incorrect keying. If one waits until the batch production edit for correction, the least expensive opportunity for correction has been bypassed -- a suspicion about keying error is most easily

---

*Some edit systems retain a flag string at the end of the record, even in the final format. Thus the record length may be increased, but the original data field locations are usually retained.

answered at keying time. Further, if one waits until the production edit, and then wishes to check against source data forms, one must find the form. NCES will be receiving data for 16,000 LEAs and 90,000 schools, making later source document recovery a time-consuming endeavor. Further, if a very difficult edit failure arises at data-entry time, the operator may be instructed to reject the record, and lay it aside for analyst intervention, before the data have even been entered into the system.

The potential of on-line data entry is impressive. There are even analysts who propose to totally eliminate batch edits, and use only on-line techniques at data-entry time.* While this extreme position is debatable on cost-effectiveness grounds, it is clear that at least partial implementation of edits at data-entry time will be valuable for CCD.

Preliminary machine edits. The topic of preliminary editing has been described extensively above, and only a few words need to be said here. The emphasis at this stage is on statistical reporting, on a state-by-state basis, of the quality of the data base; that is, on what general problems can quickly be detected for a given state's data. A reasonable goal would be to produce a preliminary edit report within seven working days after receipt of data on magnetic tape, and within 20 working days after receipt of data in hard-copy form. The specific edit checks which can be performed profitably at this stage depend on the details of the survey instrument.

Statistical reporting. Statistical reporting is not only an important component of the preliminary edit phase, but of the production edit phase as well. Once editing of the data is completed, reports should be assembled from both edit stages to provide a final picture of the quality of the data obtained

---

*Gilb, T. E., and Weinberg, G. M. Humanized Input. Cambridge, MA: Winthrop, 1977.

from each state. Analysis of these reports should dictate the kind of general follow-up and technical assistance which may be required for long-term improvement of data quality (including forms and instructions revision, better training or communication, computer systems assistance, etc.).

Automatic correction. In many cases, a data item which fails an edit should not require manual correction. Depending on the circumstances, one may choose to permit the edit program to correct data elements so as to pass the edit ("correct and log") or permit it to correct and warn the user ("correct and warn"), or permit it to suggest a correction but require the user to agree ("suggest and hold"), or permit it only to uncover the error ("notify and hold"). The particular option selected depends on what data element is involved, what kind of error has been committed, and how gross the error is. A situation in the LEA Non-Fiscal Survey where automatic correction may apply is in item 1, where, in each row of this item, column 3 is expected to be equal to the sum of columns 1 and 2. Assuming there are data in all three columns, what should be done if 1 and 2 don't sum to 3? Suppose they contained 196, 173, and 367 respectively; in this case it might be reasonable to assume that the clerk filling out the form originally made an addition error, obtaining a sum of 367 instead of 369. If this were the case, the appropriate correction would be to replace 367 with 369.* Note, however, that other explanations (and thus other corrections) are tenable. For instance, our straw-person clerk might have transcribed the value to column 1 incorrectly, recording 196 instead of 194. Thus, automatic correction in such a case might only trade one error for others. However, it was found last year when manual intervention was required prior to correction that in most cases

*Most systems would also enter a code in the appropriate flag field to indicate the field had been corrected, and would issue a notification to the user.

of addition checks, the only pragmatic correction was to adjust the sum to agree with the addends. Thus, despite the possibility that the correction might be itself in error, requiring approval of a human analyst gains almost nothing. Faced with a choice of permitting the known discrepancy to remain in the data (one kind of noise), or introducing a correction which might be in error (another kind of noise), the human analyst chooses the latter. By permitting the computer to make such replacements automatically, we save our analyst for more difficult tasks. What about the situations where such a correction is "unreasonable?" If "unreasonable" can be defined, then it can be programmed. For instance, suppose the discrepancy is quite large; in this case we might wish to correct it and draw the analyst's attention to the problem. Thus, we use the following automatic correction logic:

> If the reported sum is less than 5% deviant from the computed sum, correct and list the correction on the log;
>
> If the reported sum is more than 5% deviant from the computed sum, correct and list the correction on the "warning" report.

An even more sophisticated algorithm would be to compare the difference of the computed sum and the recorded sum to "5% or 1.0, whichever is larger" to handle the case of small absolute numbers (e.g. 7 + 2 = 9).

Of course, some errors will not easily be amenable to automatic correction (e.g., missing data). Such instances must not, however, force the data to remain in the edit system forever; there must be provision for releasing data even when they are known to be bad, and there must be provision for recording, in the data record itself, that the data are bad, and should not be used, or used only with caution (c.f.

missing value codes in standard statistical computing systems).* Such record-keeping permits the later use of imputation procedures, i.e., statistical procedures for replacing bad data values during a later, data analysis phase.

Relational/longitudinal edits. Relational editing involves the comparison of a new value for a data element to an old value(s), and examination of the amount of change the element has undergone. An example is comparing an LEA's reporting membership in the present school year to last year's response or to responses from several preceding years. The edit is performed by determining whether the extent of change from preceding values to the currently reported value is within acceptable boundaries (e.g., "+7% to -5%"). While such edits will not be possible for CCD this year, they must be considered now while the edit system is being redesigned.**

By themselves, relational edits are risky, since they permit the introduction of certain kinds of longitudinal biases into the data. If erroneous data are introduced in one year, and whatever caused the erroneous data is still operating in subsequent years, it is unlikely that the error will ever be caught. Because of this heavy dependence on the accuracy of existing data, relational edits are most often of value when used in conjunction with more traditional kinds of edits such as value or range checks and arithmetic checks. In early years, while the relational data base is still being established, paired edits (i.e., relational plus conventional) will protect one from the longitudinal bias described above; later, it may be reasonable to remove many (but not all) of the non-relational edits and to retain only those pairs that relate to highly

---

*Alternative approaches to flagging missing data are to use special values (e.g., -0.0) in the data field itself.

**The data file from last year is not complete and has not been thoroughly processed, nor will it be in time to be used in the editing of this year's data.

volatile data elements or that enhance automatic edit-check resolution.

The use of paired edits could be of special value to NCES, since it holds the potential both for reducing the number of cases that require manual processing (e.g., fail a reasonably defined conventional check) and for enhancing the system's ability to resolve some edit failures automatically. These properties can lead to substantial savings in the time and resources required by the edit process. An example may help to clarify how a relational edit, paired with a conventional edit, may be very efficient. Consider two data elements currently included in Part VI-A: student membership and FTE-teachers. Currently, these elements are edited by computing a pupil-teacher ratio and applying a range check. How would one set the acceptable range for this edit? The problem is that this ratio can vary from as high as 40 or 50 (as in some elementary schools) to as low as 1 (as in some special education classes). If the range for the edit is set this wide (i.e., $1 \leq \text{ratio} \leq 50$), it will be almost worthless, since too many bad values will slip through; conversely, if the range edit is set too narrowly, too many good values will be flagged and a substantial amount of manual intervention will be required.*

Suppose, however, that a relational edit were paired with a moderately narrow range check in the following way: if a field failed <u>both</u> checks, it would have failed edit. If, on the other hand, it passed either check (or both) it would be allowed to pass. With this arrangement, a special-education school might fail a range check but have the same pupil/teacher ratio as last year, and thus pass the paired edit. If another

---

*Even setting the range contingent on other data element values (technically a "matrix" edit), such as whether it is an elementary or secondary school, seems unsatisfactory for many cases. Legitimate ratios may vary depending on demographic and other characteristics that will never be feasible to program into the edit system. Thus, a completely suitable range edit seems beyond reach.

special-education school was opened to normal students (e.g., for mainstreaming), the pupil/teacher ratio would change radically from the preceding year, but it would then be within the (moderately narrow) definition of normal and pass the range check; since passing either of the paired checks is sufficient, it to would pass edit. A safe variant of this procedure would be to pass a field which met either of a paired set of checks, but notify the user via a warning message, so that the decision could be checked by hand. Note that the default action of the system is to pass the record (saving human effort by having the system do what the human analyst would do most of the time), but manual intervention is also facilitated by gathering for the analyst information only on those data records and fields which absolutely require his attention.

Table-driven edits. Many of the edits to be used involve the specification of numeric parameters (e.g., "within 5%" or "greater than 5 and less than 40"); based on last year's experience with CCD edits, it would be wise to design the new system so that such parameters are read from an easily accessible table, independent of the program proper. Optimally, this table would be installed on an on-line disk pack -- then, if some initial parameter values are found to be poor (as is bound to be the case, since they are only best guesses at first), they will be easy to modify.

Switch-driven edits. Just as some parameters require tuning, so does the entire edit system. Most systems use a few pieces of data in the input stream to set up basic system controls. For example, a "switch" in the input stream might be used to signal whether the data file is about to be edited for the first time (and therefore read in row one format), or has been edited before (and thus is in intermediate, edit format), or is to be output from the edit system (and thus to be written out in an output format). While such gross control switches are familiar, it might be reasonable to consider using switches far more extensively to fine-tune the edit

system. Switches could be associated with each individual edit check, and stored in an on-line table for efficient maintenance. Then specific edits might be turned on or off as appropriate. This would be especially useful for specific cases (states) where systematic edit violations are known and unavoidable. For example, suppose that the preliminary edit revealed that some detail items are always blank for a given state, and a follow-up to the state indicates that the data are simply not available. In such a case, turning off the checks on the relevant data fields would be in order, saving paper otherwise wasted with nonsense edit reports, and simplifying the task of isolating true edit violations.*

User-oriented interface. Articles and books about systems design commonly discuss the importance of the user interface -- the point at which computer output and user input cross. All that can be done in this paper is to reemphasize the point and make some suggestions; NCES should keep the issue firmly in mind when implementing these specifications.

Edit systems, in particular, are notorious for generating output measured better in pounds than lines, and for requiring user inputs (like updates, corrections, and overrides) to be in a form dictated by programming convenience. A slight increase in the design-and-programming investment to humanize the interface will result in significant savings in both cost and time during edit processing. The approach will also be effective at reducing new errors introduced as part of the edit process itself.

Messages must tell the survey analyst what is wrong, where, and provide sufficient information (at least) to allow a guess as to why. They must be arranged so that various kinds

---

*Of course, for fields with their own estimation indicator embedded in the record, the edit system should check to see if data are "not available" before proceeding with other edits.

of messages can be found without lengthy searching; messages that are defined as critical should be located separately from messages that are considered less pressing. The user should assist the system designer in figuring out how to describe what is wrong, and in selecting the kind of output that will help him determine why; the question of "where" will usually require two answers, one for the user, and a second for the edit system. We consider it reasonable to permit this particular piece of computer-oriented information to slip into the message system: a record identifier that has been uniquely defined by the edit system can lead to significant run-time efficiencies, and may even, occasionally, be the only means by which the user can distinguish records (as when two records for an LEA are received, one of which is spurious and must be deleted).

On the other side of the process, the edit system must provide the user with a convenient means of specifying the field that he wishes to change, and the value to insert. Including a parser in the edit-system-update module, so that the user can provide his input in relatively free-form, is strongly recommended. One unsophisticated scheme that can save considerable effort is to arrange to handle updates in the following form:

> record number -- comma -- field number -- comma --
> new field contents in quotes.

All spaces, except those within the quotes demarcating the new field contents, would be ignored by the parser. For example, "6457, 22, '9381'" would be interpreted by the edit system as "change field 22 in record 6457 to '9381'"; more importantly, the user can also easily interpret this form. This is clearly superior to any kind of fixed-format transaction record, and is much easier for the user to prepare. It also relieves the edit report of the burden of serving the secondary role of input form. Many systems use the same report to inform the

user of the error and to return his or her corrections to the edit system, needlessly complicating the message system and making it more difficult for the survey analyst to use.

To this point in the report the proposals have been analytical, theoretical, and general. The final two major sections deal more specifically with edit specifications at the data item level. Part VI-A is dealt with first; data-entry, preliminary edits, and production edits are recommended. A similar presentation then follows for Part VI. It should be noted at this point that many of the production edits are not original in this report but have been adapted from the edit system developed for the last cycle of CCD. They are included with the new material so as to provide a complete system view.

EDITS FOR PART VI-A -- UNIVERSE OF PUBLIC SCHOOLS

This survey involves the collection of 14 substantive and seven utility data elements on every public school in the nation. Data are collected from each of the 58 state and territorial education agencies (SEAs), which in most cases collated the data elements from their existing records.* Data are received from the states in three forms: hard-copy (the form is reproduced in Appendix A), shuttle-listing (essentially a form-facsimile that is computer-generated by NCES from data collected during the preceding VI-A cycle, and on which the respondent indicates changes only), and magnetic tape (tape format specifications are reproduced in Appendix B). The first two forms must be keyed by NCES onto machine-readable media, at which point they are equivalent in form and format to the submissions received on magnetic tape. As a result of the multiple forms of submission, three kinds of edits must be considered: on-line at data-entry time, preliminary machine, and production edits.**

Table 1 summarizes the edits proposed for the current cycle of The Public School Universe Survey. The first column of the table contains a list of the data fields, using the field name assigned by NCES (see the tape record format documentation in Appendix B). A field description is also included to assist in associating the data field with the corresponding survey item (Appendix A). Across from each data field name

---

*If a data element is not collected by a given state, it will be missing from all Part VI-A individual school records within that state's jurisdiction. In such cases, NCES may attempt to obtain the data from another source. Negotiations with the state are also held to arrange for the collection of such data elements in the future.

**A fourth kind of editing, manual/clerical editing of hard-copy, is also done, but will not be dealt with here.

Table 1

SUMMARY OF EDITS FOR PART VI-A,
PUBLIC SCHOOL UNIVERSE

	DATA ENTRY		PRELIMINARY MACHINE EDIT		PRODUCTION EDIT	
Field Name/Description	Edit Check	Action	Edit Check	Action	Edit Check	Action
1. NCES-ID/ 7-byte LEA no.	Presence, filled numeric 1st 2 bytes must be valid OE state code	Reject on failure Reject on failure	Presence Filled numeric 1st 2 bytes	Count failures Frequency distrib.	Valid LEA id number	Reject on failure; write to special hold files and request user action
2. SCHNO/ 5-byte numeric NCES school no.	Presence, filled numeric Else blank if field #21 = "N"	Reject on failure (check validity of field #21) Write message to operator to set aside forms with "NEW" schools.	Presence Filled numeric Presence	Count failures Crosstab with field #21	Presence, filled numeric Else blank if field #21 = "N"	If absent & school is "NEW", assign no. from available pool. If present & school is "closed", retire number If absent & school is not new or is closed, reject record, writing it to special hold file, & write message to user requesting action.
3. SEA-ID/ SEA's id code for LEA (20-byte alphanumeric)		Left-justify	Presence	Count failures		Left-justify
4. SYS-NAME/ Name of LEA (30-byte alpha-numeric)	Presence	Reject on failure (if not available enter "UNKNOWN" & confirm.) Left-justify	Presence	Count failures	Presence	Auto-correct on failure to "UNKNOWN" & write warning message allowing override update. Left-justify
5. SEASCHID/ SEA's id code for school (20-byte alpha-numeric)		Left-justify	Presence	Count failures		Left-justify
6. SCH-NAME/ School name (30-byte alpha-numeric)	Presence	Reject on failure (if not available, enter "UNKNOWN" & confirm) Left-justify	Presence	Count failures Frequency count on "UNKNOWN" vs blank vs all other.	Presence	Auto-correct on failure to "UNKNOWN" & write warning message allowing override update. Left-justify
7. STADDEST/ 1-byte estimation indicator for field #8	Legal code (blank or "N") Consistency check (field #8)	Auto-correct on failure, field #8 dominant, notify operator.	Legal code Consistency check	Frequency count Crosstab with presence check on field #8	Legal code (blank or "N") Consistency check (field #8)	Auto-correct on failure, field #8 dominant; notification to log.
8. ST-ADDR/ Street address (25-byte alpha-numeric)	Presence	Left-justify Consistency check against field #8. Request operator confirmation.	Presence	Count failures Crosstab with field #7.	Presence	Left-justify (see field #7); write warning message to log.
Compound check on fields #2, #5, #6, #3	Simultaneous absence	Reject of fields 2, 5, 8 are empty & 6 is empty or "UNKNOWN"			Simultaneous absence	Reject if fields 2, 5, & 8 are empty & 6 is "UNKNOWN"; write record to special hold file & request user action.
9. CITY-EST/ 1-byte indicator for field #10	Legal code (blank or "N") Consistency check (field #10)	Auto-correct on failure, field #8 dominant; notify operator.	Legal code Consistency check	Frequency count Crosstab with presence check on field #10	Legal Code (blank or "N") Consistency check (field #10)	Auto-correct on failure, field #10 dominant; notification to log.

Table 1 (cont'd.)

SUMMARY OF EDITS FOR PART VI-A,
PUBLIC SCHOOL UNIVERSE

Field Name/Description	DATA ENTRY		PRELIMINARY MACHINE EDIT		PRODUCTION EDIT	
	Edit Check	Action	Edit Check	Action	Edit Check	Action
10. CITY/ City name, 13-byte alphanumeric	Presence	Left-justify	Presence	Count failures	Presence	Left-justify
		Consistency check against field #9; request operator confirmation		Crosstab with field #9		(see field #9); write warning message to log.
11. ST-ABBRV/ 2-byte state abbreviation	Presence	Auto-correct on failure using 1st 2 bytes of field #1 (OE code) & notify operator.	Presence	Count failures	Presence	Auto-correct on failure; generate value using 1st 2 bytes of field #1 (OE code)
	Legal code if present (U.S. Postal Service)	Reject on failure.	Legal code	Frequency Distrib.	Legal code if present (U.S. Postal Service)	Auto-correct as above; write warning message allowing override/update (in case of error in field #1)
12. ZIPCDEST/ 1-byte indicator for field #13	Legal code (blank or "N")	Auto-correct on failure, field #13 dominant	Legal code	Frequency Distrib.	Legal code (blank or "N")	Auto-correct on failure, field #13 dominant; notification to log.
	Consistency check (field #13)		Consistency check	Crosstab with presence check on field #13	Consistency check (field #13)	
13. ZIP-CD-5/ 5-byte numeric	Presence	Consistency check against field #12	Presence	Count failures	Presence	If filled numeric & consistent with field #12, pass.
	Filled numeric	On failure, reject or update field #12 to "N", blank field #13	Filled numeric	Crosstab with field #12	Filled numeric	If blank & consistent with #12, pass.
						If filled or blank & inconsistent with #12, update #12, notify log, & pass.
						If non-blank & not filled, force #12 to "N", #13 to blank, & write warning message allowing override/update.
14. SCH-TYPE 1-byte numeric	Legal code ("1" to "7")	Reject on failure (operator could correct to blank)	Legal code	Frequency Distrib.	Legal code ("1" to "7")	Auto-correct to blank, warning message allowing override/update.
15. GRD-SPAN LO/ Lower limit of grade span, 2-byte alphanumeric (see also #16)	Legal code ("PK", "KG", "01" to "12", "UC" blank)	Reject on failure	Legal code	Frequency Distrib.	Legal code ("PK", "KG", "01" to "12", "UC", blank)	Auto-correct to blank, write warning message allowing override/update
16. GRD-SPAN HI/ Upper limit of grade span, 2-byte alphanumeric	Legal code (see #15)	Reject on failure	Legal code	Frequency Distrib.	Legal code (see #15)	Auto-correct to "UC" if #15 = "UC" else blank & write warning message allowing update/override
	Consistency check #16=blank if #15 blank, #16="UC" if #15="UC", else #16 ≥ #15	Reject or forced auto-correct	Consistency check	Crosstab with field #15 to verify #16 ≥ #15 or both blank		Auto-correct & write warning message

Table 1 (cont'd.)

Summary of Edits for Part VI-A,
Public School Universe

Field Name/Description	DATA ENTRY		PRELIMINARY MACHINE EDIT		PRODUCTION EDIT	
	Edit Check	Action	Edit Check	Action	Edit Check	Action
Consistency check on fields #14, #15, #16	Consistency check #14 #15 & #16 1 "UC", "PK", "KG", "01" thru "08" 2 Any 3 or 4 of "03" thru "09" 3 "UC", "07" thru "12" 4 Any 5 Any 6 "07" thru "12" 7 Any		Consistency check #14 #15 & #16 1 "UC", "PK", "KG", "01" thru "08" 2 Any 3 or 4 of "03" thru "09" 3 "UC", "07" thru "12" 4 Any 5 Any 6 "07" thru "12" 7 Any	Crosstab	Consistency check #14 #15 & #16 1 "UC", "PK", "KG", "01" thru "08" 2 Any 3 or 4 of "03" thru "09" 3 "UC", "07" thru "12" 4 Any 5 Any 6 "07" thru "12" 7 Any	Auto-correct #15 & #16 to blank; write warning message allowing update/ override
17. TCH-EST/ 1-byte indicator for field #18	Legal code ("*", "β", "N", or blank)	Auto-correct to N if field #18 blank, else auto-correct to "β" & warn (ambiguity "β" vs "*")	Legal code	Frequency Distrib.	Legal code ("*", "β", "N", or blank)	Auto-correct to N if field #18 blank, else auto-correct to "β" & write warning message
	Consistency check		Consistency check	Crosstab with presence check on field #18	Consistency check	
18. TEACHFTE/ FTE of classroom teachers, 4-byte numeric fixed point no. in F4.1 format	Presence	Consistency check against #17	Presence	Frequency Distrib.	Presence	Consistency check against #17 on failure, write warning; auto-correct to blank; auto-correct #17 to N
	Justified, filled numeric	Operator must enter no. right-justified, leading zeroes; reject on failure	Range check	Mean, Mauian, 5 highest values, 5 lowest values	Justified, filled numeric	
	(NOTE: this could be set up as a split field on entry to force decimal)					
19. MEMB-EST/ 1-byte indicator for field #20	Legal code ("*", "β", "N", or blank)	Auto-correct to N if field #20 blank, else auto-correct to "β" & warn (ambiguity "β" vs "*")	Legal code	Frequency Distrib.	Legal code ("*", "β", "N", or blank)	Auto-correct to N if field #20 blank, else auto-correct to "β" & write warning message
	Consistency check		Consistency check	Crosstab with presence check on field #20	Consistency check	
20. MEMBRSHP/ 4-byte numeric pupil membership	Presence, numeric	Consistency check against #19	Presence	Frequency Distrib.	Presence	Consistency check against #19 on failure; write warning; auto-correct to blank; auto-correct #19 to N
		Right-justify	Range check	Mean, Median, 5 highest values, 5 lowest values	Justified, filled numeric	
Consistency check against fields #18, #20/ #20 ÷ 18 if both present pupil/teacher ratio	If #18 & #20 are present If school type ≠ 5 #20 ÷ #18 is ≥ 12.0 ≤ 35.0; Else #20 ÷ #18 ≥ 3.0 ≤ 20.0	On failure, request verification of #18 & #20	Range check	Compute mean, median, 5 highest values, 5 lowest values, by school type	If #18 & #20 are present If school type ≠ 5 #20 ÷ #18 is ≥ 12.0 ≤ 35.0; Else, omit check	Write warning message on failure to allow update/ override; leave fields #18 & #20 unchanged.
21. NEWCLOSD/ 1-byte alphanumeric field to indicate school status (new, closed, existing)	Legal code ("N", "C", blank)	Reject on failure	Legal code	Frequency Distrib.	Legal code ("N", "C", blank)	Auto-correct to blank if field #2 has legal entry else reject; write to special hold file & request user action
	Consistency check	See field #2	Consistency check	Crosstab with presence on field #2		

and description are three pairs of columns, one pair corresponding to each kind of edit treated herein: on-line at data-entry, preliminary machine, and production. The left-hand column of each pair describes the condition to be screened for (e.g., "presence" -- is something present in the field?), while the right-hand column generally indicates the action to be taken if the edit is failed (e.g., "if nothing is present in the field -- if it is blank -- the record should be rejected").

Because terminology is not well standardized in this area, it will be useful to discuss the proposed edits in some detail. In order to organize the presentation, we will treat each phase of editing as an entity, going down the pairs of columns of Table 1, one at a time. Later we will discuss some implications that are only apparent when the edits are examined across phases.

Data Entry Edits

The first data field contains the NCES numeric identification number assigned (for the most part) uniquely to every LEA in the country. It is a fundamental identifier in NCES's data collection system, and may be cross-referenced to other data bases, including those of the Bureau of the Census. It is a critical data item, and is absolutely required. The edit criteria at data entry are stringent -- the field must have data ("presence"), and the data must be, precisely, a 7-digit number, since all LEA numbers are of this form. Further, the first two digits (left-most) must correspond to a valid Office of Education (OE) code for one of the 50 states, the District of Columbia, or 7 territories. This means that the first two digits must be a number between 10 and 69. If the record fails to meet any of these criteria, the data-entry system must reject the record. In practice this means that the key operator is notified that the record has been rejected, but the record remains on the CRT screen. The operator would then visually verify the keyed record to determine if a keying error had been made. If so, the operator would correct the

field "on-line" and resubmit it to the system. If no keying error had been made, and the entry in the raw form was in fact not legal, then the operator would be instructed to set the form aside, and to pass it on at the appropriate time to an NCES analyst. (This analyst would then have to resolve the problem before it was ever entered into the system.) Having set the form aside, the operator would proceed to key subsequent records.

Another edit could be included for this field, but its value is questionable in a real-time system. The system could check the entire 7-digit number against a table of legal LEA codes, but this table contains approximately 16 000 entries, organized into only 58 groups, and thus the check would be fairly costly in computer resources. On the other hand, this check might be very cost-effective in a batch edit environment (see Production Edit column for field #1).

Field #2 is the NCES school number, which is to be a unique identifier for each school in a given state (i.e., it must be combined with the two-digit state code to be unique nationally). Edits are specified for this field in Table 1, since numbers were assigned for many states during the 1977-78 tryout of this survey. However, SAGE has recommended new permanent numbers be assigned this year, in which case this edit would be skipped at data-entry.* If the field is to be edited, the check would be similar to that for the first field: Most legal entries should consist of 5-digit numeric entries; in addition, the field could legitimately contain all zeroes or blanks, but only if field #21 (NEWCLOSD") indicates that it is a new school. Any time field #2 contained other than numbers or blanks, or any time it failed the numeric-presence check and field #21 did not indicate a new school, the record

*Fingerman, P. W. Letter report to Mr. Warren Hughes, Institutional Surveys Branch, Division of Elementary and Secondary Education Statistics, NCES. 7 March 1979.

would be rejected. In addition, it is recommended that when field #2 contains all zeroes or blanks and field #21 indicates the school is new (i.e., legal entries signalling a new school), the record be validated.* The number of new schools is relatively small, and their verification is reasonable compared to the risk of fouling the identification scheme for such an important longitudinal data base.*

The third field contains the state education agency's (SEA) code for the LEA, if any. It is an optional field maintained for the convenience of the states, should they request the data base in the future, or should they request a shuttle-list for future responses. The notation, "left-justify" in the action column means that leading blanks, if any, should be stripped off by the data-entry program.

The fourth field is reserved for the name of the LEA. Table 1 indicates a presence check: reject the record if there is nothing in this field, and perhaps the data entry operator to enter a place holder, "UNKNOWN", to fill the field. This field is also to be left-justified.

The fifth field is meant to contain the SEA's code (if any) for the school. It is treated just as the third field is.

The sixth field is for school name, and an entry is required. Thus, if present, the operator keys the name, and the program left-justifies it. If no name is present, the program substitutes "UNKNOWN" and asks the operator to confirm by checking the raw form.

The next pair of fields are linked: field #7 is an indicator field for field #8, street address of the school. Field #7 must be blank unless the street address is not available,

---

*This field could be afforded considerable additional protection were a check-digit added. The logic for handling check-digits is automatic in most commercial data entry systems. This topic is examined again in the context of the production edit system.

in which case it should contain "N" (for "not available"). This is set up primarily for tape submissions of data, to verify that the address is not available in state records, as distinct from accidentally omitted or missing. Field #8 is left-justified, and should be checked to see if something has been entered. If so, field #7 is expected to be a blank. If there is a street address and the indicator is an "N", field #7 should be automatically corrected to a blank (i.e., field #8 is "dominant" over field #7), and the key operator should be notified. This notification would provide an opportunity for the operator to correct the situation if the entry in field #8 were accidental, and field #7 correctly contained an "N". If there is no entry under street address, and field #7 contains a blank, the system converts this blank and requests operator confirmation.

The next edit involves a check for simultaneous absence of all identification information regarding the school. Thus, if a record has no NCES school number, no state school number, no school name (or "UNKNOWN"), and no address, then it is not an acceptable record. In other words, some information is required which distinguishes the school from others in the same LEA.

The ninth and tenth fields are paired in the same way as #7 and #8. Number 9 is an indicator for #10, city name. The edit checks are analogous to those for #7 and #8 respectively.

Field #11 is a two-character data item containing the U.S. Postal Service's abbreviation for the state. A check for presence is made and, if the field is blank, the system supplies a state abbreviation using the two-digit OE state code from field #1. A check is also made to verify that the entry is a "legal code," i.e., a legitimate abbreviation for one of the 58 reporting units. If the entry is not legal, the record is rejected. The operator may then correct the entry if possible, or put it aside for later handling by an NCES analyst.

The next two fields, #12 and #13, are again an indicator-data field pair, and the consistency of the relationship must be checked as between fields #7 and #8. In addition, field #13 is to contain a zip code, and thus the five digits must be filled.

Field #14 is a one-digit item indicating type of school. It can take on seven legal values ("1" to "7"); if it is not one of these the record is rejected. Fields #15 and #16 contain data on grade span, and are subject to the same kind of legal code check. Field #16 is further required to be consistent with #15. The next entry in Table 1 is a matrix edit for fields #14, #15, and #16, designed to check the consistency of the information in these three fields.

Field #17 is an indicator field for field #18, which is to contain data on the number of full-time equivalent (FTE) classroom teachers in the school. Field #17 is checked for a legal code, and for consistency with #18 (and automatically corrected to correspond to #18, notifying the operator if such correction is necessary). According to the instructions (Appendix B), the number entered in #18 is to the nearest tenth, but no decimal point is keyed. This introduces the possibility of an incorrect entry due to mispositioning the value in the field. For example, if the correct value were 101.1, and it were miskeyed as " 101", it would be interpreted as 10.1; if the correct value were 10.0 and it were miskeyed as "100 ", it could be interpreted at 100.0.* Thus the edit requirement that the operator enter the value right-justified and zero-filled is recommended to reduce keying errors. An alternative would be to set up the data entry program to accept this field as two entries, an integer part and a fraction part, and to require the fraction part be entered (i.e., a required field). In any event, the point is to prevent positional errors in entering this item.

---

*Keying " 101" means, literally, keying blank-one-zero-one.

Fields #19 and #20 are another indicator pair; once again they must be checked for consistency. In addition, #20 is a numeric field, which must contain only numbers or leading blanks.

If both fields 18 and 20 contain data, a further edit is possible: a matrix range check on pupil/teacher ratio. As noted above, this check is not optimal, but is often the best alternative when no relational checks are available (see below).

The last field is used to indicate whether the school is new, has closed during the past year, or continues to operate. Three codes are legal, and are checked for; the consistency check between this field and #2 (NCES school number) has previously been discussed. It may be possible in the future, in cases of disagreement between this field and #2, to verify the school characteristics (e.g., name) against values from the preceding year to resolve the discrepancy (see relational edits below).

Three final points should be emphasized. First, data entry operators should be clearly instructed on what to do when a record fails an edit. Generally, they should be told to check to make sure no keying error has been made. If such an error is discovered, the data-entry program should permit them to make a correction and then to edit the record again. If the error is contained in the source document, the document should be put aside for review by subject matter analysts (e.g., the survey sponsor and staff). Second, the edit program must be flexible. Edits must be adjustable, especially range-checks (see table-driven edits, above). The program should permit edits to be over-ridden or "soft-failed" when they are advisory in nature, or when a proportion of correct records is expected to fail a given range-check edit (e.g., pupil/teacher ratios). If these also fail a production edit, you will know that the failure was <u>not</u> due to a keying error (since it failed and was checked at entry time). Finally, the data-entry program should allow the edits to be turned off. One example of when this is useful is when input is to be keyed <u>and</u> verified. Since edit

checks are performed at entry time, it is a waste of on-line resources to continue to perform edit checks again during verification. Only records which are changed during verification require re-editing, and the program should be set to turn edits back on automatically in this case.

The final point has to do with relational/longitudinal edits and on-line data entry. Data base access methods and state-of-the-art computer technology make such edits feasible in an interactive environment, although somewhat expensive. If NCES expects to continue to receive a substantial proportion of data in hard-copy, then a move to add such edits will probably be justified. Whether performed at entry or held until batch, the setting of relational edit parameters, i.e., how much deviation from last year's value should be tolerated before the comparison is considered suspicious, depends on the lability of the data item. Later this year SAGE plans to conduct empirical studies on the longitudinal behavior of some CCD data items in order to lay the groundwork for relational editing in the next cycle.*

Preliminary Machine Edits

The preliminary machine edits presented here were developed with two factors in mind: first, since they are the first edits for data submitted in machine-readable form, they must be sensitive to many of the same errors which drove the design of the data entry edits. Second, the primary purpose of the preliminary edit is to discover statistically systematic edit problems which might be fixed programmatically as, for example, when a state frequently but consistently uses an alternative code value (e.g., "K " for "KG"). Such errors can often be fixed prior to the production edit with a simple reformatting or transformation program.

---

*As mentioned previously, such edits are not possible this year.

For these reasons, preliminary machine edits will often be found to be less thorough but partially redundant with data entry edits. Nevertheless, it is recommended that keyed data be passed through the preliminary machine edit program. Often a bias may thus be revealed that was not apparent looking at records one at a time during data entry. For example, suppose that a state accidentally ignored the "NEWCLOSD" field, and failed to report schools that were opened or closed since the last survey. The data-entry system would catch the new schools (no school number in field #2, and no indication of new in the "NEWCLOSD" field) via a consistency check; such records would be rejected, and the operator would set aside the form. However, the absence of closed schools would not be caught by the entry system, and the operator is not likely to notice either. A frequency distribution on the "NEWCLOSD" field produced by the preliminary machine edit program, however, would quickly reveal the complete absence of closings (and, if run before the analyst reviewed and corrected the set-aside new schools, the absence of openings). If this were not sufficiently suspicious, a quick comparison with the number of school status changes reported last year by this state might settle the matter.

The specific preliminary machine edits proposed have been coded into an SAS program by SAGE. They are also listed in the middle pair of columns in Table 1. No field-by-field treatment is necessary here since these edit checks are essentially a subset of those described above for data entry. However, the "actions" shown are entirely different, corresponding to the statistical purpose of these edits. The actions of this edit are various kinds of counting, and the outputs are generally frequency distributions (either one-way, or cross tabulations). In addition, ranges of numeric items are determined. Finally, alphabetic characters in numeric fields are scanned for, and a detailed report is provided for (up to) an arbitrary number of such errors (the default is 50, mostly to save paper). If more than the default number of illegal

numeric conversions are found, they will be counted, and a field-by-field summary printed.

Despite the thoroughness of the screening provided by this program, the full output will rarely exceed 20 pages, making it practical to determine the overall quality of a newly submitted magnetic tape almost immediately upon receipt. Follow-up when problems are revealed is facilitated since the analyst at the state level will still have the project fresh in mind. Also, the rapidity of the screening provides a maximum amount of time for writing special-purpose programs to fix any systematic problems uncovered.

While no relational or longitudinal edits are included in the program code prepared this year, it would be simple to add such checks using SAS. In fact, experimental versions of this kind of code will be derived during the empirical studies of longitudinal editing proposed for SAGE later this year.

Production Edits

In many respects, the production edit system is an extension of the data-entry edit system, executed on large batches of records in one run. Many of the same edits are included; in fact, the version of the production system recommended for this year extends only slightly the protection afforded by data entry checks. The system is necessary this year nevertheless, and will become increasingly valuable in the future. First, it is required since a substantial portion of the data will be received on magnetic tape, thereby not being amenable to NCES screening at data entry. The production edits also provide an opportunity to check entries that have been corrected subsequent to initial data entry (e.g., during verification). Perhaps most importantly, more sophisticated methods of automatic correction of suspicious and erroneous data items can be employed, because of the nature of the interaction between the computer edit system and the subject matter analyst. The production system should be permitted to make a great number of changes in flagged data automatically, programmed

according to the available experience and provided with a series of "best guesses". This is done in full knowledge that all of its decisions are subject to review and that its message system is arranged to call our attention to those changes about which we are least confident.

The batch system is specified to add a few edit checks that were considered optional at data entry time, primarily because of the additional computing resources required. Resources are often available in background tasks at lower cost than in an interactive foreground environment. Finally, this system will be the future locus for an extensive set of relational/longitudinal edits.

Because of the similarity between the data entry and production edits, only salient differences will be discussed in detail here. The reader is referred again to Table 1, and particularly to the first (data entry) and third (production) pairs of columns on the table. The first field represents a situation in which a more stringent check is proposed for the production system. At data entry time a check on the pattern of the NCES LEA identifier was recommended, guaranteeing the presence of a seven-digit number which (at least) could be a legal LEA code. In the production edit it is proposed that the LEA identifier be verified against a list of legal values. A further, optional check (not in Table 1) would be to verify both the LEA identifier and the LEA name against a list of legal entries.*

---

* Special methods are required for matching records on alphanumeric fields like LEA name if one is to avoid too many false non-match occurrences. One technique is to strip out vowels, blanks, and special characters from two fields to be compared before testing for a match. However, this much protection may not be necessary at present. The clerical screening of hard-copy survey material affords one opportunity of checking the validity of this identifier, and manual spot-checks on machine-readable submissions may be more efficient than including such an edit for magnetic tape files on a routine basis. Such spot-checks could be included as part of the preliminary edit program, implemented by having that program print a sample list of LEA numbers from the files which could then be validated by clerical staff.

The second field, school number, is perhaps the most troublesome. While LEA numbers have been in use for some time, and are widely familiar to both respondents and data users, permanent school numbers are an entirely new form of identification which may require some time before they are well established. If the SAGE's recommendation is followed, and new permanent numbers are assigned this year, little editing is required on this field (e.g., a check for presence and numeric fill). In future cycles a variety of relational edits would be imposed on this field, including a check of reported number and reported name for consistency against the preceding cycle. No parallel edit is feasible this year if new numbers are not assigned.*

The production edit program has an additional responsibility with regard to the school-number field: it controls the status of such numbers, assigning new ones to new schools, and retiring numbers assigned to schools which close. For this function it must access and maintain a school-number file which indicates the active, available, and retired numbers on a state-by-state basis. The program should also maintain an audit trail of all activity against this file for historical purposes.** The reader's attention is directed to Table 1 for a description of routine control procedures.

---

*Despite the fact that the survey content has been frozen through 1981, serious consideration should be given to modifying the hard-copy, shuttle-list, and tape formats, in order to extend the length of the school number field by one byte. This byte would be used to hold a modulus-11 check-digit for the school number field, increasing the integrity of this field considerably. This change in format could be made next year, using school numbers generated this year which contain check-digits from the start. Only the shuttle-list states would be impacted upon at all, and even those only slightly.

**Provision must be made for manual modifications to this system of school numbering, either via a special entry point in the production edit program, or through the use of a coordinated file maintenance program.

Mention was made above regarding the interaction between the operation of the production edit system and the subject matter analyst. Several entries in Table 1 illustrate this. One such kind of interaction is indicated for fields which indicate, under the Action column, "reject on failure; write to special hold file and request user action." These fatal errors are associated with so-called required fields, fields whose content is considered absolutely vital (e.g., #1, #2). The action entry indicates the writing of the input record to a special file rather than to the normal, intermediate edit file. The record is held there until the subject matter analyst (the "user") takes some action. Generally, this message is used to signal an edit failure that can only be remedied by human action. A message is written to a special output file (see below) indicating the serious nature of the problem, and as much information about the record as is available (e.g., dump of all relevant data items and identifiers).

Other examples of edit/analyst interaction are illustrated by the entries "write notification to the log" (e.g., field #7), "write warning message", and "write warning message allowing correction/update" (fields #8, #11). The edit program uses several different output files for messages: fatal errors ("reject on failure" above) will be written to one file, and the analyst must respond. Less critical errors (those associated with warning messages) are written to a second file, and will usually report that some default correction has been employed that should be reviewed for appropriateness. Finally, some messages are only notifications, for audit purposes, of fairly safe actions that the edit system has taken. These should be carefully examined early in the life of the system to ensure that they are, indeed, safe; later they will only require occasional scanning, and perhaps statistical treatment to analyze error behavior among respondents.

As indicated above, it is the production edit system that is most likely to execute most of the relational edits when

they are introduced. In addition to the value of such edits in checking identification fields like LEA number, their application to non-identification numeric fields is also important. The use of a relational check paired with a matrix range check for pupil/teacher ratio was discussed earlier in this report and need not be repeated. Planned work by SAGE will lay the groundwork for using such edits in future cycles. They are mentioned again here, in the context of proposed edit specifications, primarily to remind computer analysts who translate these specifications to keep such modifications in mind during system design.

EDITS FOR PART VI -- LEA NON-FISCAL SURVEY

This survey collects data on a large number of quantitative variables for each of the approximately 16,000 LEAs in the country. These data, consisting of tabulations of staff by sex and function, of students by grade and type of school, and of several other miscellaneous items, are collected from each of the 58 SEAs, just as the school-level data (Part VI-A) are. As is the case with the school universe survey, data are for the most part collated from existing records by each SEA, and are returned to NCES on hard-copy (the form is reproduced in Appendix C) or magnetic tape (tape format specifications are reproduced in Appendix D). Thus, three forms of edit are again called for: on-line edit at data entry time to guard against keying error, preliminary machine editing for the detection of systematic errors, and production editing of batched records prior to the final release of the data for analysis.

The survey data can be subdivided into two major classes, identification and utility items (e.g., LEA name, address, "new-closed" indicator) and enumerated data items (items 1 through 9 of the form). The edits recommended for the former class are very similar to those performed for analogous components of the Part VI-A survey, and are summarized in Table 2. Four of these items are edited using checks identical to those recommended for Part VI-A; three (street address, city, and zip code) differ only in that the Part VI survey includes no indicator field to accompany them, and so no consistency check is required in Table 2. The last item in both surveys, "NEWCLOSD", is subjected to different edits in Part VI. In this survey the item is used to indicate the consolidation ("closing") or division ("opening") of LEAs themselves, and is subjected only to a check for legal code. However, because of the relative rarity of a change in status for an LEA, a

Table 2

Summary of Edits for Part VI, LEA Non-Fiscal Report:
Identification Fields

l Name/Description	DATA ENTRY		PRELIMINARY MACHINE EDIT		PRODUCTION EDIT	
	Edit Check	Action	Edit Check	Action	Edit Check	Action
KTS-ID/ -byte numeric LEA number	Presence, filled numeric	Reject on failure	Presence, filled numeric	Count failures (blanks, zeroes)	Valid LEA id number	Reject on failure; wri to special hold file request user action
	lst 2 bytes must be valid OE state code	Reject on failure	lst 2 bytes	Frequency Distrib.		
YS-NAME/ lame of LEA (30-byte alpha-numeric)	Presence	Rejection failure (if not available, enter "UNKNOWN" & confirm)	Presence	Count failures ("UNKNOWN", blank)	Presence	Auto-correct on failur to "UNKNOWN" & write warning message allow override/update
		Left-justify				Left-justify
T-ADDR/ treet address (13-byte alpha-numeric)		Left-justify	Presence	Count failures (blank)		Left-justify
ITY/ ity name (13-byte alpha-numeric)		Left-justify	Presence	Count failures (blank)		Left-justify
T-ABBRV/ tate abbreviation (2-bytes)	Presence	Auto-correct on failure using lst 2 bytes of field #1 (OE code) & notify operator	Presence	Count failures (blank)	Presence	Auto-correct on failur generate using lst 2 bytes of field #1 (OE code)
	Legal code (U.S. Postal Service)	Reject on failure	Legal code	Frequency Distrib.	Legal code (U.S. Postal Service)	Auto-correct as above; write warning message allowing override/upd
IP-CODE/ -byte numeric	Filled numeric or blank	Reject on failure	Presence, filled numeric	Count failures (blanks, zeroes)	Presence, filled numeric or blank	Auto-correct to blank failure; write notifi tion to log.
EA-ID EA id for LEA (20-byte alpha-numeric)		Left-justify	Presence	Count failures (blank)		Left-justify
NCLOSD/ EA status (1-byte alpha-numeric)	Legal code (blank, "N", "C")	Reject on failure	Legal code	Frequency Distrib.	Legal code (blank, "N", "C")	Write warning message user on any non-blank entry

warning message should be written whenever the field is non-blank; note also that new LEAs without NCES identification numbers are written onto a special hold file during the production edit phase so that an identification number can be assigned.

Enumerated data items in the Non-Fiscal Survey include 123 fields. Associated with each of these numeric fields is an indicator/estimator field which may take on one of four codes: blank or zero when the number in the accompanying data field is an active value, "*" when the datum in the field is an estimate, and "N" when the datum in the field is not available. In this last case, the data field itself is to contain all zeroes (an NCES specification). Thus, the first step in editing these numeric data items is to check the consistency of each data field with its associated indicator/estimator field. At data entry time, the following rules should be applied:

- If the indicator/estimator is blank, zero, or "*" and the data field contains a non-zero numeric value, allow the field to pass edit;
- If the indicator/estimator is blank, zero, or "*" and the data field contains the quantity zero, request operator confirmation in order to pass the field (thus trapping for missing data when no "N" was keyed into the indicator/estimator field);
- If the indicator/estimator is blank, zero, or "*" and the data field contains non-numeric data or is blank, reject the record (the operator, of course, may correct and resubmit it if the problem is a keying error);
- If the indicator/estimator contains an "N" and the data field contains the quantity zero, allow the field to pass edit;
- If the indicator/estimator contains an"N" and the data field is other than zero-filled, reject the record;
- If the indicator/estimator contains any code other than a blank, zero, "*", or "N", reject the record.

Finally, for any record that passes this edit, the data-entry system should right-justify and zero-fill the data field.

The preliminary machine edit for these fields should be a simple cross-tabulation of indicator/estimator code by the three relevant field conditions (numeric-zero, numeric-non-zero, or non-numeric) for each of the 123 items. The production edit program should apply the same standards as those applied at data entry.*

Once the data have passed this initial set of consistency checks, attention must turn to the consistency among data fields, (e.g., the checking of arithmetic, the reasonableness of the values). This is the primary means of verifying that the numbers themselves are accurate. Some very powerful consistency checks are available for the first item on the survey (staff); somewhat less powerful checks are available for the remaining items 2 through 9. Because these latter checks are easier to describe, we shall deal with them first.

The second item on the survey is an enumeration of membership by grade and type of school (elementary and secondary). Assuming that data are available, the sum of the elementary school membership fields (2A, 2B, 2C, 2D, 2E, 2F, 2G, 2H, 2Ia, 2Ja, and 20a) should equal the elementary-total field (2Pa). Upon data entry, the following procedure should be applied: if the sum is equal to the total field, pass the item; if the sum is within ± 5% of the total field but not equal to it, auto-correct the total field to equal the computed sum, and ask the operator for confirmation; if the sum deviates more than ± 5%, reject the record.** The same procedure should be

---

*If the preliminary edit reveals that some consistent problem regarding the use of these codes exists in a state's magnetic tape submission, one might turn off these consistency edits instead of writing a special transformation/translation program. The real purpose of the indicator/estimator fields is the detection of estimated data, but this purpose should probably be held subordinate to the goal of processing the data in a timely fashion.

**A better standard than "± 5%" would be "± 5% or within ± 2, whichever is larger." This permits small totals with small errors to be corrected automatically. The discussion above is limited to "± 5%" only as a matter of convenience.

applied during production edits, except that when a 5% deviation is auto-corrected, a notification should be written to the log, and when a deviation in excess of 5% is found, auto-correction should be applied (total set equal to computed sum) and a warning message sent to the user. All of these auto-corrections must be applied with caution: if the edit is performed when the individual grade enrollments are missing (indicators are "N"), the data fields will sum to zero (since they are always zero-filled). Thus, this check-sum-and-auto-correct procedure must only be applied when all requisite data are present.

The same kind of data entry and production edits are called for in dealing with the secondary school membership fields (sum of fields 2Ib, 2Jb, 2K, 2L, 2M, 2N, and 2Ob should equal the total field, 2Pb). In addition, entries for seventh-grade elementary and seventh-grade secondary (2Ia and 2Ib) should be checked against one another; only one should have data (for most school districts). If both have non-zero values, the data entry program should request operator confirmation, and the production edit program should write a notification to the log. The same kind of check should be applied to the two eighth-grade fields (2Ja and 2Jb).

The preliminary machine edit program should print the five highest and lowest entries and the mean for each field in item 2. In addition, the elementary and secondary school membership sums should be computed, the appropriate total fields subtracted, and the difference evaluated for each LEA: the five highest and five lowest differences and the associated totals should be printed, along with the mean differences between computed sums and reported totals. Finally, frequency distributions on the associated indicator/estimator fields should be printed.

Items 3, 4, and 5 are all similar in content, and should be subjected to the same arithmetic check: providing the data are available, the sum of the first two columns (e.g., 3A and 3B) should be equal to the total column (e.g., 3C). When it

is not, but is within ± 5%, the data entry program should set the total to the computed sum and request operator confirmation, while the production edit program should make the same correction and notify the log; if the deviation is greater than 5%, the data entry program should reject the record, while the production program should auto-correct and write a warning message.

In addition to this common arithmetic check, one additional check should be applied to item 3 alone. Item 3C should be non-zero if students are reported in the twelfth grade (2Pb) and both data values are present. Write a warning message on failure (data entry and production edit).

The preliminary machine edit program should compute the five highest and five lowest values and the mean for: the data fields, the computed sum for each item, and the differences between the reported and computed totals. The preliminary edit program should also produce frequency distributions for each of the relevant indicator/estimator fields, and a crosstab presence check between fields 3C and 2Pb.

For items 6, 7, 8, and 9 no edit checks of any value at the individual record level are possible, since there is little information internal to the survey against which to check them; powerful edits on these fields will have to await the next survey cycle, when relational edits become available.* One step should be taken this year to try and deal at least with extreme instances of error: the preliminary machine edit program should produce a frequency distribution of the relevant indicator/estimator fields, as well as the 20 highest and lowest values, the mean, and the standard deviation for each of the data fields.

Returning to item 1, much more powerful edit checks may be applied to these data, and more sophisticated methods of auto-correction for errors are available. This results from

---

*Item 8 could be checked against the school universe file this year, but this would probably not be cost-justified.

the relative redundancy built into the item, the number of interlocked detail, sub-total and total cells that mutually constrain the values which each may take on. The approach still depends on the requirement that a reported sum of two or more fields should be equal to the computed sum; the factor which distinguishes item 1 is that every data field is involved in at least two such arithmetic comparisons. This acts as a constraint, and guarantees the identification and accurate correction of any record with one error among the item 1 fields, and a high probability of accurate detection and correction of records with two or more errors. Before laying out the complete system, a simple example will be used to illustrate the method.

Figure 2a depicts a portion of item 1 with illustrative data. It includes three of the thirty rows of item 1, corresponding to part M. Each row is divided (as is the rest of item 1 -- see Appendix C) into three columns, "male", "female", and "total". The illustrative data add correctly, both across rows (job total and grand total for the four detail fields) and down columns (sex total, and grand total for the four detail fields). For the examples below, assume that these are the "true" data, which might be erroneous when the data record is ultimately received by NCES. Consider now Figure 2b. When the addition checks are performed on this array of data, we find that the third row does not add up, marked by the arrow to the right of the table, and the second column does not add up, indicated by the arrow beneath the table. One cell entry has been transcribed in error, and it can easily be seen that this cell is in the third row, second column. Further, the arrows "pointing" to the bad row and the bad column intersect at the cell containing the bad datum. Moreover, only one value can replace the bad datum and satisfy both of the addition check failures: replacing the "7" in the critical cell with the correct value, "6", satisfies both additions. Figure 2c illustrates this phenomenon for a similar situation, where one cell contains an error, and this cell is located in the single row and the single column that fail to add properly.

a) correct:

Assignment/function	Male (a)	Female (b)	Total (c)
M. Aides			
1. Instructional aides	1	2	3
2. Other aides ...	3	4	7
3. TOTAL (M 1 and 2)	4	6	10

b) one error:

Assignment/function	Male (a)	Female (b)	Total (c)
M. Aides			
1. Instructional aides	1	2	3
2. Other aides	3	4	7
3. TOTAL (M 1 and 2)	4	7 ←	10
		↑	

c) one error:

Assignment/function	Male (a)	Female (b)	Total (c)
M. Aides			
1. Instructional aides	1	3 ←	3
2. Other aides	3	4	7
3. TOTAL (M 1 and 2)	4	6	10
		↑	

d) two errors:

Assignment/function	Male (a)	Female (b)	Total (c)
M. Aides			
1. Instructional aides	1	2	4 ←
2. Other aides	3	4	8 ←
3. TOTAL (M 1 and 2)	4	6	10
			↑

e) two errors:

Assignment/function	Male (a)	Female (b)	Total (c)
M. Aides			
1. Instructional aides	1	5	3 ←
2. Other aides	6	4	7 ←
3. TOTAL (M 1 and 2)	4	6	10
	↑	↑	

Figure 2. Auto-Correction Method for Item 1 - Part VI

In fact, it can be proven that, for any table with r rows and c columns where one of the rows is generated by summing the other rows and one of the columns is generated by summing the other columns, a change in any single value in the table will lead to exactly two addition check failures, one for a row and one for a column, and that the cell at the intersection of this row and column must be the cell that is inconsistent. Thus, for any table with exactly one error it is possible to locate and correct the error with perfect confidence. Conversely, any table that has more than one row or more than one column that do not add up correctly contains at least two bad cell entries.

What about tables with exactly two errors? Here the situation is more complex. If the two errors are located in the same row or column, the cells can be identified and unique corrections are computable (Figure 2d); the key to identifying this situation is noting that there are exactly two rows and no more than one column, or two columns and no more than one row that fail to add properly. Two errors that are not in the same row or column are not uniquely identifiable or correctable; such situations may be distinguished from those above by the combination of two rows and two columns which fail the addition check. With more than two errors, not even the kind of error (e.g., one cell, two cells in same row or column, two cells in different rows and columns) may be identified.

The point is that we have the capability not only to identify errors (by addition check edits) but also, in a substantial number of cases, to correct them automatically with a high degree of confidence in the accuracy of the correction. How then might this system be applied to all of item 1? Begin by noting that item 1 may be partitioned into four sub-sections or partitions which themselves, meet the criteria spelled out above (i.e., r by c tables with a sum-column and a sum-row). Figure 3 illustrates these partitions, outlined with heavy lines. One further partition in Figure 3 includes the "total" rows from

1. Full-time equivalent number of persons employed by this agency during the payroll period including October 1, 1978.

REPORT TO THE NEAREST TENTH

	Assignment/function	Male (a)	Female (b)	Total (c)
*	A. Superintendents			
*	B. Other officials/administrators			
	C. Principals			
	1. Elementary			
	2. Secondary			
	3. Unclassified			
	D. Assistant principals			
	1. Elementary			
	2. Secondary			
	3. Unclassified			
*	E. Total of principals & asst. principals			
*	F. Curriculum specialists			
*	G. Library/media specialists			
*	H. Psychological personnel			
	I. Classroom teachers			
	1. Prekindergarten			
	2. Kindergarten			
	3. Other elementary			
	4. Secondary			
	5. Unclassified			
*	6. TOTAL *(I 1 thru 5)*			
*	J. Other teachers, e.g., radio/TV, etc.			
	K. Guidance & counseling personnel			
	1. Elementary			
	2. Secondary			
	3. Unclassified			
*	4. TOTAL *(K 1 thru 3)*			
*	L. Other professional personnel			
	M. Aides			
	1. Instructional aides			
	2. Other aides			
*	3. TOTAL *(M 1 and 2)*			
*	N. Office/clerical personnel			
*	O. Other nonprofessional personnel			
*	P. Total, all personnel			

Figure 3. Partitions of Item 1, Part VI.

each of the four partitions noted above plus all of the rows not included in one of the four other partitions; these are marked on the figure by the asterisks to the left of each included row. Thus, there are five partitions of item 1 which may be subjected to the kind of edit treatment described above.

Each partition should be checked in the following manner:

- If the partition meets the row and column addition checks, proceed to the next partition;
- If the addition checks indicate one cell in error, compute the size of the error (computed sum to reported sum) -- if it is within ± 5%, auto-correct the cell entry, note the change on the log, and proceed to the next partition -- if it is more than 5%, auto-correct the cell, write a warning message, and proceed;
- If the addition checks indicate two cells in the same row or column are in error, proceed as with one error;
- If the addition checks do not uniquely isolate multiple errors examine the topmost row which does not add properly -- if the total for this row can be adjusted by ± 5% to bring the row into balance, do so and recheck the table -- if it still does not check, correct any uniquely identifiable cell error (based on the second check) or correct the topmost row total that fails to check by up to ± 5% and recheck -- if the partition meets the check, proceed to the next partition -- if not, reject the record -- in any event, write a warning message to the user, and set an edit system flag -- in the event any other checks are generated during subsequent editing of item 1 on that record, the record should be rejected and held for manual checking, and a warning message should be written for the user.*

The order in which these partitions should be checked is "inside-out": inside partitions A, B, C, D, followed by the outside partition marked by asterisks (Figure 3). Fields which are changed by this edit should have their indicator/estimator fields changed to "X" to record the event.

---

*Otherwise the program could spend a week making up spurious corrections for one really bad record.

This kind of edit consumes computer resources fairly heavily, and may not be feasible in many on-line data entry systems; it should be included as a component of the production edit system. Two other edit checks should also be applied to item 1, as part of both data entry and production edits:

- If total membership (item 2Pa plus 2Pb) exceeds 10,000, items 1Fc, 1Gc, 1Hc, 1Jc, 1K4c, 1Lc, 1M3c, 1Nc, and 1Oc should all have non-zero entries (if data is available);
- Field 1I6c (total teachers) should be greater than 1Pc (total personnel) minus 1I6c (i.e., there should be more teachers than any other kind of personnel).

Records that fail these edits at data entry should be rejected; those that fail during production edit should be rejected and written to a special hold file for later correction. If they are, indeed, correct (but unusual), the production edit must be arranged to allow the user to force the record past these edits.

While these edits all are implementable this year during the current cycle of Part VI, there is the potential, as there was for Part VI-A, for implementing relational/longitudinal checks next year. With the exception of item 1, such checks would significantly improve the power of error detection; in the case of item 1 the improvement is not as clear, given the powerful edit technique that will already be in place. It is recommended that the actual performance of the two alternative approaches be compared empirically this summer (using what data are available from last year's tryout of Part VI for the relational checks) to determine whether both are justified when used simultaneously, or, if not, which should be preferred for future use.

APPENDIX A

This appendix contains a copy of the survey form used by NCES for the 1978-79 cycle of CCD Part VI-A, the Universe of Public Schools survey.

PARTMENT OF HEALTH, EDUCATION, AND WELFARE
EDUCATION DIVISION
WASHINGTON, D.C. 20202

ORE OF DATA, PART VI-A. UNIVERSE OF PUBLIC SCHOOLS

LOCAL EDUCATION AGENCY IDENTIFICATION
NCES SEA
NAME OF LOCAL EDUCATION AGENCY

FORM APPROVED
O.M.B. No. 51R-1227

is authorized by law (20 U.S.C. 1221e-1). While you are not required to respond, your cooperation is make the results of this survey comprehensive, accurate, and timely.

TIONS. List the schools operated by each local education agency in the state along with the data for each sted. If continuation sheets are needed, please make sure that the next higher line number is inserted on h continuation sheet. Line numbers should be in sequence from 0001 to the highest number needed for a agency. The number of schools listed on this part should equal the number of schools entered for Item ES Form 2393-2.

Codes to be used:

School Type
1 = elementary
2 = middle
3 = secondary
4 = combined elementary/secondary
5 = special education for handicapped
6 = vocational/technical
7 = alternative

Grade Span — Use 01, 02, 03, . . . 12 for numbered grades. Use PK for prekindergarten. Use KG for Kindergarten. If the school is graded, enter the appropriate grade designations for the lowest and highest grades offered. If the school is ungraded, enter UG-UG in the grade span column.

(continue on the reverse)

L IO	LINE NO.	NAME OF SCHOOL	STREET ADDRESS	CITY	ZIP CODE	SCHOOL TYPE	GRADE SPAN		FTE OF TEACHERS*	OCTOBER 1 MEMBERSHIP
							LOW	HIGH		
	0001									
	0002									
	0003									
	0004									
	0005									
	0006									
	0007									
	0008									
	0009									
	0010									
	0011									
	0012									
	0013									
	0014									
	0015									

2A, 9/78 *(FM Control No. 70)*

*REPORT TO THE NEAREST TENTH

APPENDIX B

This appendix contains keying instructions and the tape record layout prepared by NCES for CCD Part VI-A, 1978-79.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page 212-01 4
Subsection(s) .01
Date Documented 10/27/78
☒ New ☐ Revision
Change Notice #

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX KEYTAPE

212.01 KEYING INSTRUCTIONS

CHARGE # SYSTEM ID: CCD PAGE 1 OF 2
JOB TITLE DATE: 10/27/78
SOURCE DOCUMENT PART VI-A. UNIVERSE OF PUBLIC SCHOOLS FORM # 2393-2A
COMMENTS

CARD NUM.	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		NCES-ID	1	7	7	N	KEY 7 NUMERICS
		SCHNO	8	12	5	N	SEE NOTE #1. SPACE FILL
		SEA-ID (SEA's ID FOR THE LEA)	13	32	20	A/N	LEFT JUSTIFY, SPACE FILL TO RIGHT. SPACE FILL IF MISSING.
		SYS-NAME (AGENCY)	33	62	30	A/N	" " " " "
		SEASCHID (SEA'S SCHOOL IDENT)	63	82	20	A/N	" " " " "
		SCH-NAME	83	112	30	A/N	" " " " "
		STADDEST	113	113	1	A/N	KEY "N" IF NOT AVAILABLE. KEY SPACE IF PROVIDED.
		ST-ADDR	114	138	25	A/N	LEFT JUSTIFY, SPACE FILL TO RIGHT. SPACE FILL IF MISSING.
		CITY-EST	139	139	1	A/N	KEY "N" IF NOT AVAILABLE. KEY SPACE IF PROVIDED.
		CITY	140	152	13	A/N	LEFT JUSTIFY, SPACE FILL TO RIGHT. SPACE FILL IF MISSING.
		BLANK	153	154	2	A/N	LEAVE BLANK (SPACE FILL)
		ST-ABBRV	155	156	2	A	SEE NOTE #4. SPACE FILL.
		BLANK	157	158	2	A/N	LEAVE BLANK (SPACE FILL)
		ZIPCDEST	159	159	1	A/N	KEY "N" IF NOT AVAILABLE. KEY SPACE IF PROVIDED.
		ZIP-CD-5	160	164	5	N	KEY 5 DIGIT ZIP CODE.
		BLANK	165	168	4	A/N	CONTINUE ZIP CODE IF IT IS 9 DIGITS LONG. OTHERWISE SPACE FILL.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page 212-01 5
Subsection(s) .01
Date Documented 10/27/78
X New Revision
Change Notice #

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX KEY TAPE

212.01 KEYING INSTRUCTIONS

CHARGE # SYSTEM ID: CCD PAGE 2 OF 2
JOB TITLE DATE: 10/27/78
SOURCE DOCUMENT PART VI-A. UNIVERSE OF PUBLIC SCHOOLS FORM # 2393-2A
COMMENTS

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		SCH-TYPE	169	169	1	N	KEY A "1-7".
		GRD-SPAN LO	170	171	2	A/N	KEY "PK" IF PREKINDERGARTEN KEY "KG" IF KINDERGARTEN
		GRD-SPAN HI	172	173	2	A/N	KEY "UC" IF UNCLASSIFIED OTHERWISE KEY "01-09, 10-12
		TCH-EST	174	174	1	A/N	SEE NOTE #2.
		TEACHFTE	175	178	4	N	SEE NOTE #3.
		MEMB-EST	179	179	1	A/N	SEE NOTE #2.
		MEMBRSHP	180	183	4	N	
		NEWCLOSD	184	184	1	A/N	SPACE = SCHOOL IS OPEN. N = NEW SCHOOL. "C" = CLOSED.
		BLANK	185	200	16	A/N	SPACE FILL.

NOTE #1. SCHNO (I.E. SCHOOL NUMBER) WILL BE COMPUTER GENERATED BY NCES. NUMBERING WILL BEGIN WITH 0001 WITHIN EACH STATE. SEA'S DESIRING TO PERFORM THIS TASK MAY DO SO BUT SHOULD NOTIFY MR. HUGHES OF NCES IN WRITING.

NOTE #2. THE ESTIMATION INDICATOR FIELDS CONTAIN THE SUFFIX "EST" IN THEIR NAMES. ACCEPTABLE CODES ARE: "*" = ESTIMATED DATA. ZERO OR SPACE = TRUE DATA. "N" = DATA NOT AVAILABLE.

NOTE # 3. THE FTE OF TEACHERS WILL BE ROUNDED TO THE NEAREST TENTH. THE DECIMAL WILL BE UNDERSTOOD BUT NOT PRESENT ON THE TAPE.

NOTE # 4. NCES WILL COMPUTER GENERATE THE P.O. STATE ABBREVIATIONS IF THEY ARE NOT PROVIDED BY THE SEA.

APPENDIX C

This appendix contains a copy of the survey form used by NCES for the 1978-79 cycle of CCD Part VI, the Local Education Agency Nonfiscal survey.

DEPARTMENT OF HEALTH, EDUCATION, & WELFARE
EDUCATION DIVISION
NATIONAL CENTER FOR EDUCATION STATISTICS

FORM APPROVED
O.M.B. NO. 51R-1227

DUE DATE: January 15, 1979

COMMON CORE OF DATA - PART VI. LOCAL EDUCATION AGENCY NONFISCAL REPORT

This report is authorized by law (20 U.S.C. 1221e-1). While you are not required to respond, your cooperation is needed to make the results of this survey comprehensive, accurate, and timely.

ID numbers	Name of agency
NCES	Street address
SEA	City, State, ZIP

1. Full-time equivalent number of persons employed by this agency during the payroll period including October 1, 1978.

REPORT TO THE NEAREST TENTH

Assignment/function	Male (a)	Female (b)	Total (c)
A. Superintendents			
B. Other officials/administrators			
C. Principals			
1. Elementary			
2. Secondary			
3. Unclassified			
D. Assistant principals			
1. Elementary			
2. Secondary			
3. Unclassified			
E. Total of principals & asst. principals			
F. Curriculum specialists			
G. Library/media specialists			
H. Psychological personnel			

NCES FORM 2393-2, 9/78 *(FM Control No. 76)*

1. Full-time equivalent number of persons employed by this agency during the payroll period including October 1, 1978. (continued)

Assignment/function	Male (a)	Female (b)	Total (c)
I. Classroom teachers			
1. Prekindergarten			
2. Kindergarten			
3. Other elementary			
4. Secondary			
5. Unclassified			
6. **TOTAL** *(I 1 thru 5)*			
J. Other teachers, e.g., radio/TV, etc.			
K. Guidance & counseling personnel			
1. Elementary			
2. Secondary			
3. Unclassified			
4. **TOTAL** *(K 1 thru 3)*			
L. Other professional personnel			
M. Aides			
1. Instructional aides			
2. Other aides			
3. **TOTAL** *(M 1 and 2)*			
N. Office/clerical personnel			
O. Other nonprofessional personnel			
P. Total, all personnel			

II. Number of pupils in membership on October 1, 1978, or nearest date thereto when a fall membership count is taken.

If as of date is not October 1, please specify the date: _____

Grade level	Elementary (a)	Grade level	Elementary (a)	Secondary (b)
A. Prekindergarten		I. Seventh		
B. Kindergarten		J. Eighth		
C. First		K. Ninth		
D. Second		L. Tenth		
E. Third		M. Eleventh		
F. Fourth		N. Twelfth		
G. Fifth		O. Unclassified		
H. Sixth		P. TOTALS *(A thru O)*		

III. Number of 12th grade graduates from the regular day school program (including summer session) during the 1977-78 school year.

Male	Female	Total

IV. Number of pupils scheduled to be transported at public expense on or about October 1, 1978

Public School	Private School	Total

V. Number of vehicles used to transport pupils owned wholly or jointly by the agency on or about October 1, 1978
[Large = more than 15 passenger]
[Small = less than 16 passenger]

Large	Small	Total

VI. Total area enclosed within the agency's boundries in square miles .. _____

VII. Number of members of the board of education ..... _____

VIII. Number of schools operated by this agency on October 1, 1978 ..... _____

IX. Number of scheduled days in the regular school term when pupils are expected to be in attendance ..... _____

SEE PAGE 4 FOR SPECIAL INSTRUCTIONS

SPECIAL INSTRUCTIONS TO ACCOMPANY NCES FORM 2393-2

1. Report the full-time equivalent number of persons employeed to the nearest tenth.

2. If personnel within selected assignment categories cannot be reported by level, report the appropriate totals only.

3. If the number of principals and assistant pricipals cannot be reported separately, enter "N.A." on page 1, lines C and D, and enter the total number of principals and assistant principals on the same page, line E.

4. If the number of aides cannot be reported separately by type, report the total number of aides on Line M.3.

5. These forms should be returned to:

NCES/DESES/ISB
Federal Office Building No. 6
400 Maryland Avenue, SW
Washington, D.C. 20202

APPENDIX D

This appendix contains keying instructions and the tape record layout prepared by NCES for CCD Part VI, 1978-79.

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX . KEYTAPE

212.01 KEYING INSTRUCTIONS

CHARGE # ___ SYSTEM ID: CCD PAGE 1 OF 17

JOB TITLE ___ DATE: 10/26/78

SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT FORM # 2393-2 (5-1-78)

COMMENTS ___

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO?	A/N	INSTRUCTIONS
		NCES-ID	1	7	7	N	KEY 7 NUMERICS
		SYS-NAME	8	37	30	A/N	SPACE FILL IF MISSING. LEFT JUSTIFY & SPACE FILL.
		ST-ADDR	38	62	25	A/N	" " " " "
		CITY	63	75	13	A/N	SEE P.O. PUBLICATION 59 FOR ABBRV. OF NAMES OVER 13 CHARACTERS LONG.
		BLANK	76	77	2	A/N	KEY TWO SPACES
		ST-ABBRV	78	79	2	A	SEE ZIP CODE DIRECTORY FOR P.O. STATE ABBRV.
		BLANK	80	81	2	A/N	KEY TWO SPACES
		ZIP-CODE	82	86	5	N	KEY 5 DIGIT ZIP CODE
		BLANK	87	90	4	A/N	CONTINUE ZIP CODE IF IT IS 9 DIGITS LONG.
		SEA-ID	91	110	20	A/N	LEFT JUSTIFY & SPACE FILL
		1AAEST	111	111	1	A/N	NOTE 3.
	1AA	MALE SUPERINTENDENTS	112	117	6	N	NOTES 1 & 2.
		1ABEST	118	118	1	A/N	NOTE 3.
	1AB	FEMALE SUPERINTENDENTS	119	124	6	N	NOTES 1 & 2.
		1ACEST	125	125	1	A/N	NOTE 3.
	1AC	TOTAL SUPERINTENDENTS	126	131	6	N	NOTES 1 & 2.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page 212-01 15
Subsection(s) .01
Date Documented 10/26/78
X New ☐ Revision
Change Notice #

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX KEYTAPE

212.01 KEYING INSTRUCTIONS

CHARGE # SYSTEM ID: CCD PAGE 2 OF 17
JOB TITLE DATE: 10/26/78
SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT FORM # 2393-2 (5-1-78)
COMMENTS

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		1BAEST	132	132	1	A/N	NOTE 3.
	1BA	OTHER MALE OFFICIALS/ ADMINISTRATORS	133	138	6	N	NOTES 1 & 2.
		1BBEST	139	139	1	A/N	NOTE 3.
	1BB	OTHER FEMALE OFFICIALS/ ADMINISTRATORS	140	145	6	N	NOTES 1 & 2.
		1BCEST	146	146	1	A/N	NOTE 3.
	1BC	TOTAL OTHER OFFICIALS/ ADMINISTRATORS	147	152	6	N	NOTES 1 & 2
		1C1AEST	153	153	1	A/N	NOTE 3.
	1C1A	ELEMENTARY MALE PRINCIPALS	154	159	6	N	NOTES 1 & 2.
		1C1BEST	160	160	1	A/N	NOTE 3.
	1C1B	ELEMENTARY FEMALE PRINCIPALS	161	166	6	N	NOTES 1 & 2.
		1C1CEST	167	167	1	A/N	NOTE 3.
	1C1C	TOTAL ELEMENTARY PRINCIPALS	168	173	6	N	NOTES 1 & 2.
		1C2AEST	174	174	1	A/N	NOTE 3.
	1C2A	SECONDARY MALE PRINCIPALS	175	180	6	N	NOTES 1 & 2.
		1C2BEST	181	181	1	A/N	NOTE 3.
	1C2B	SECONDARY FEMALE PRINCIPALS	182	187	6	N	NOTES 1 & 2.

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page 212-01 16

Subsection(s) .01

Date Documented 10/26/78

X New ___ Revision

Change Notice # ___

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX KEYTAPE

212.01 KEYING INSTRUCTIONS

CHARGE # ___ SYSTEM ID: CCD PAGE 3 OF 17

JOB TITLE ___

SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT DATE 10/26/78

COMMENTS ___ FORM # 2393-2 (5-1-78)

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		1C2CEST	188	188	1	A/N	NOTE 3.
	1C2C	TOTAL SECONDARY PRINCIPALS	189	194	6	N	NOTES 1 & 2.
		1C3AEST	195	195	1	A/N	NOTE 3.
	1C3A	UNCLASSIFIED MALE PRINCIPALS	196	201	6	N	NOTES 1 & 2.
		1C3BEST	202	202	1	A/N	NOTE 3.
	1C3B	UNCLASSIFIED FEMALE PRINCIPALS	203	208	6	N	NOTES 1 & 2.
		1C3CEST	209	209	1	A/N	NOTE 3.
	1C3C	TOTAL UNCLASSIFIED PRINCIPALS	210	215	6	N	NOTES 1 & 2.
		1D1AEST	216	216	1	A/N	NOTE 3.
	1D1A	ELEMENTARY MALE ASSISTANT PRINCIPALS	217	222	6	N	NOTES 1 & 2.
		1D1BEST	223	223	1	A/N	NOTE 3.
	1D1B	ELEMENTARY FEMALE ASSISTANT PRINCIPALS	224	229	6	N	NOTES 1 & 2.
		1D1CEST	230	230	1	A/N	NOTE 3.
	1D1C	TOTAL ELEMENTARY ASSISTANT/PRINCIPALS	231	236	6	N	NOTES 1 & 2.
		1D2AEST	237	237	1	A/N	NOTE 3.
	1D2A	SECONDARY MALE ASSISTANT PRINCIPALS	238	243	6	N	NOTES 1 & 2.

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Subsection(s) .01

Date Documented 10/26/78

X New ___ Revision

Change Notice # ___

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX KEYTAPE

212.01 KEYING INSTRUCTIONS

CHARGE # ___ SYSTEM ID: CCD PAGE 4 OF 17

JOB TITLE ___ DATE 10/26/78

SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT FORM # 2393-2 (5-1-78)

COMMENTS ___

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		1D2BEST	244	244	1	A/N	NOTE 3.
	1D2B	SECONDARY FEMALE ASSISTANT PRINCIPALS	245	250	6	N	NOTES 1 & 2.
		1D2CEST	251	251	1	A/N	NOTE 3.
	1D2C	TOTAL SECONDARY ASSISTANT PRINCIPALS	252	257	6	N	NOTES 1 & 2.
		1D3AEST	258	258	1	A/N	NOTE 3.
	1D3A	UNCLASSIFIED MALE ASSISTANT PRINCIPALS	259	264	6	N	NOTES 1 & 2.
		1D3BEST	265	265	1	A/N	NOTE 3.
	1D3B	UNCLASSIFIED FEMALE ASSISTANT PRINCIPALS	266	271	6	N	NOTES 1 & 2.
		1D3CEST	272	272	1	A/N	NOTE 3.
	1D3C	TOTAL UNCLASSIFIED ASSISTANT PRINCIPALS	273	278	6	N	NOTES 1 & 2.
		1EAEST	279	279	1	A/N	NOTE 3.
	1EA	TOTAL MALE PRINCIPALS & ASSISTANT PRINCIPALS	280	285	6	N	NOTES 1 & 2.
		1EBEST	286	286	1	A/N	NOTE 3.
	1EB	TOTAL FEMALE PRINCIPALS & ASSISTANT PRINCIPALS	287	292	6	N	NOTES 1 & 2.
		1ECEST	293	293	1	A/N	NOTE 3.
	1EC	TOTAL PRINCIPALS & ASSISTANT PRINCIPALS	294	299	6	N	NOTES 1 & 2.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page 212-01 18
Subsection(s) .01
Date Documented 10/26/78
X New ___ Revision
Change Notice # ___

SYSTEM TITLE:	SYSTEM ID:
CORE DATA	CCD

212.XX

212.01 KEYING INSTRUCTIONS

CHARGE # ___ SYSTEM ID: CCD PAGE 5 OF 17
JOB TITLE ___ DATE: 10/26/78
SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT FORM # 2393-2 (5-1-78)
COMMENTS ___

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		1FAEST	300	300	1	A/N	NOTE 3.
	1FA	MALE CURRICULUM SPECIALISTS	301	306	6	N	NOTES 1 & 2.
		1FBEST	307	307	1	A/N	NOTE 3.
	1FB	FEMALE CURRICULUM SPECIALISTS	308	313	6	N	NOTES 1 & 2.
		1FCEST	314	314	1	A/N	NOTE 3.
	1FC	TOTAL CURRICULUM SPECIALISTS	315	320	6	N	NOTES 1 & 2.
		1GAEST	321	321	1	A/N	NOTE 3.
	1GA	MALE LIBRARY / MEDIA SPECIALISTS	322	327	6	N	NOTES 1 & 2.
		1GBEST	328	328	1	A/N	NOTE 3.
	1GB	FEMALE LIBRARY/ MEDIA SPECIALISTS	329	334	6	N	NOTES 1 & 2.
		1GCEST	335	335	1	A/N	NOTE 3.
	1GC	TOTAL LIBRARY/ MEDIA SPECIALISTS	336	341	6	N	NOTES 1 & 2.
		1HAEST	342	342	1	A/N	NOTE 3.
	1HA	MALE PSYCOLOGICAL PERSONNEL	343	348	6	N	NOTES 1 & 2.
		1HBEST	349	349	1	A/N	NOTE 3.
	1HB	FEMALE PSYCOLOGICAL PERSONNEL	350	355	6	N	NOTES 1 & 2.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page 212-01 19
Subsection(s) .01
Date Documented 10/26/78
[X] New [ ] Revision
Change Notice # _____

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX KEYTAPE

212.01 KEYING INSTRUCTIONS

CHARGE # _____ SYSTEM ID: CCD PAGE 6 OF 17

JOB TITLE _____ DATE: 10/26/78

SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT FORM # 2393-2 (5-1-78)

COMMENTS _____

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		1HCEST	356	356	1	A/N	NOTE 3.
	1HC	TOTAL PSYCOLOGICAL PERSONNEL	357	362	6	N	NOTES 1 & 2.
		1I1AEST	363	363	1	A/N	NOTE 3.
	1I1A	PREKINDERGARTEN MALE CLASSROOM TEACHERS	364	369	6	N	NOTES 1 & 2.
		1I1BEST	370	370	1	A/N	NOTE 3.
	1I1B	PREKINDERGARTEN FEMALE CLASSROOM TEACHERS	371	376	6	N	NOTES 1 & 2.
		1I1CEST	377	377	1	A/N	NOTE 3.
	1I1C	TOTAL PREKINDERGARTEN CLASSROOM TEACHERS	378	383	6	N	NOTES 1 & 2.
		1I2AEST	384	384	1	A/N	NOTE 3.
	1I2A	KINDERGARTEN MALE CLASSROOM TEACHERS	385	390	6	N	NOTES 1 & 2.
		1I2BEST	391	391	1	A/N	NOTE 3.
	1I2B	KINDERGARTEN FEMALE CLASSROOM TEACHERS	392	397	6	N	NOTES 1 & 2.
		1I2CEST	398	398	1	A/N	NOTE 3.
	1I2C	TOTAL KINDERGARTEN CLASSROOM TEACHERS	399	404	6	N	NOTES 1 & 2.
		1I3AEST	405	405	1	A/N	NOTE 3.
	1I3A	OTHER ELEM. MALE CLASSROOM TEACHERS	406	411	6	N	NOTES 1 & 2.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page 212-01 20
Subsection(s) .01
Date Documented 10/26/78
X New Revision
Change Notice #

SYSTEM TITLE:	SYSTEM ID:
	CCD

212.XX KEYTAPE

212.01 KEYING INSTRUCTIONS

SYSTEM ID: CCD PAGE 7 OF 17

JOB TITLE DATE: 10/26/78

SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT FORM # 2393-2 (5-1-

COMMENTS

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		1I3BEST	412	412	1	A/N	NOTE 3.
	I13B	OTHER ELEM. FEMALE CLASSROOM TEACHERS	413	418	6	N	NOTES 1 & 2.
		1I3CEST	419	419	1	A/N	NOTE 3.
	I13C	TOTAL OTHER ELEM. CLASSROOM TEACHERS	420	425	6	N	NOTES 1 & 2.
		1I4AEST	426	426	1	A/N	NOTE 3.
	I14A	MALE SECONDARY CLASSROOM TEACHERS.	427	432	6	N	NOTES 1 & 2.
		1I4BEST	433	433	1	A/N	NOTE 3.
	I14B	FEMALE SECONDARY CLASSROOM TEACHERS	434	439	6	N	NOTES 1 & 2.
		1I4CEST	440	440	1	A/N	NOTE 3.
	I14C	TOTAL SECONDARY CLASSROOM TEACHERS	441	446	6	N	NOTES 1 & 2.
		1I5AEST	447	447	1	A/N	NOTE 3.
	I15A	UNCLASSIFIED MALE CLASSROOM TEACHERS	448	453	6	N	NOTES 1 & 2.
		1I5BEST	454	454	1	A/N	NOTE 3.
	I15B	UNCLASSIFIED FEMALE CLASSROOM TEACHERS	455	460	6	N	NOTES 1 & 2.
		1I5CEST	461	461	1	A/N	NOTE 3.
	I15C	TOTAL UNCLASSIFIED CLASSROOM TEACHERS	462	467	6	N	NOTES 1 & 2.

OFFICE OF EDUCATION

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page [illegible]

Subsection(s) .01

Date Documented 10/26/78

X New ___ Revision

Change Notice # ___

SYSTEM TITLE:	SYSTEM ID:
	CCD

212.XX KEYSTAT.

212.01 KEYING INSTRUCTIONS

CHARGE # ___ SYSTEM ID: CCD PAGE 8 OF 17

JOB TITLE ___ DATE: 10/26/78

SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT FORM # 2392-2 (5-1-78)

COMMENTS ___

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		1I6AEST	468	468	1	A/N	NOTE 3.
	1I6A	TOTAL MALE CLASSROOM TEACHERS	469	474	6	N	NOTES 1 & 2.
		1I6BEST	475	475	1	A/N	NOTE 3.
	1I6B	TOTAL FEMALE CLASSROOM TEACHERS	476	481	6	N	NOTES 1 & 2.
		1I6CEST	482	482	1	A/N	NOTE 3.
	1I6C	TOTAL CLASSROOM TEACHERS.	483	488	6	N	NOTES 1 & 2.
		1JAEST	489	489	1	A/N	NOTE 3.
	1JA	TOTAL OTHER MALE CLASSROOM TEACHERS	490	495	6	N	NOTES 1 & 2.
		1JBEST	496	496	1	A/N	NOTE 3.
	1JB	TOTAL OTHER FEMALE CLASSROOM TEACHERS.	497	502	6	N	NOTES 1 & 2.
		1JCEST	503	503	1	A/N	NOTE 3.
	1JC	TOTAL OTHER CLASSROOM TEACHERS.	504	509	6	N	NOTES 1 & 2.
		1K1AEST	510	510	1	A/N	NOTE 3.
	1K1A	MALE ELEMENTARY GUIDANCE/COUNSELLORS	511	516	6	N	NOTES 1 & 2.
		1K1BEST	517	517	1	A/N	NOTE 3.
	1K1B	FEMALE ELEMENTARY GUIDANCE/COUNSELLORS	518	523	6	N	NOTES 1 & 2.

Subsection(s) .01

Date Documented 10/26/78

☒ New ☐ Revision

Change Notice #

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX KEYING:

212.01 KEYING: INSTRUCTIONS

CHARGE # SYSTEM ID: CCD PAGE 9 OF 17

JOB TITLE DATE: 10/26/78

SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT FORM # 2393-2 (5-1-78)

COMMENTS

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		1K1CEST	524	524	1	A/N	NOTE 3.
	1K1C	TOTAL ELEMENTARY GUIDANCE/COUNSELLORS	525	530	6	N	NOTES 1 & 2.
		1K2AEST	531	531	1	A/N	NOTE 3.
	1K1A	MALE SECONDARY GUIDANCE/COUNSELLORS	532	537	6	N	NOTES 1 & 2.
		1K2BEST	538	538	1	A/N	NOTE 3.
	1K2B	FEMALE SECONDARY GUIDANCE/COUNSELLORS	539	544	6	N	NOTES 1 & 2.
		1K2CEST	545	545	1	A/N	NOTE 3.
	1K2C	TOTAL SECONDARY GUIDANCE/COUNSELLORS	546	551	6	N	NOTES 1 & 2.
		1K3AEST	552	552	1	A/N	NOTE 3.
	1K3A	UNCLASSIFIED MALE GUIDANCE/COUNSELLORS	553	558	6	N	NOTES 1 & 2.
		1K3BEST	559	559	1	A/N	NOTE 3.
	1K3B	UNCLASSIFIED FEMALE GUIDANCE/COUNSELLORS	560	565	6	N	NOTES 1 & 2.
		1K3CEST	566	566	1	A/N	NOTE 3.
	1K3C	TOTAL UNCLASSIFIED GUIDANCE/COUNSELLORS	567	572	6	N	NOTES 1 & 2.
		1K4AEST	573	573	1	A/N	NOTE 3.
	1K4A	TOTAL MALE GUIDANCE/COUNSELLORS	574	579	6	N	NOTES 1 & 2.

212.01 KEYING INSTRUCTIONS

CHARGE #_____ SYSTEM ID: CCD PAGE 10 OF 17

JOB TITLE_____ DATE: 10/26/78

SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT FORM # 2393-2 (5-1-78)

COMMENTS_____

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		1K4BEST	590	590	1	A/N	NOTE 3.
	1K4B	TOTAL FEMALE GUIDANCE/COUNSELLORS	591	596	6	N	NOTES 1 & 2.
		1K4CEST	587	587	1	A/N	NOTE 3.
	1K4C	TOTAL GUIDANCE/ COUNSELING PERSONNEL	588	593	6	N	NOTES 1 & 2.
		1LAEST	594	594	1	A/N	NOTE 3.
	1LA	OTHER MALE PROFESSIONAL PERSONNEL	595	600	6	N	NOTES 1 & 2.
		1LBEST	601	601	1	A/N	NOTE 3.
	1LB	OTHER FEMALE PROFESSIONAL PERSONNEL	602	607	6	N	NOTES 1 & 2.
		1LCEST	608	608	1	A/N	NOTE 3.
	1LC	TOTAL OTHER PROFESSIONAL PERSONNEL	609	614	6	N	NOTES 1 & 2.
		1M1AEST	615	615	1	A/N	NOTE 3.
	1M1A	MALE INSTRUCTIONAL AIDES	616	621	6	N	NOTES 1 & 2.
		1M1BEST	622	622	1	A/N	NOTE 3.
	1M1B	FEMALE INSTRUCTIONAL AIDES	623	628	6	N	NOTES 1 & 2.
		1M1CEST	629	629	1	A/N	NOTE 3.
	1M1C	TOTAL INSTRUCTIONAL AIDES	630	635	6	N	NOTES 1 & 2.

OFFICE OF EDUCATION

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page

Subsection(s) .01

Date Documented 10/26/78

[X] New [ ] Revision

Change Notice #

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX · KEY[illegible]

212.01 KEY[illegible] INSTRUCTIONS

CHARGE # SYSTEM ID: CCD PAGE 11 OF 17

JOB TITLE DATE 10/26/78

SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT FORM # 2393-2 (5-1-78)

COMMENTS

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		1M2AEST	636	636	1	A/N	NOTE 3.
	1M2A	OTHER MALE AIDES	637	642	6	N	NOTES 1 & 2.
		1M2BEST	642	643	1	A/N	NOTE 3.
	1M2B	OTHER FEMALE AIDES	644	649	6	N	NOTES 1 & 2.
		1M2CEST	650	650	1	A/N	NOTE 3.
	1M2C	TOTAL OTHER AIDES	651	656	6	N	NOTES 1 & 2.
		1M3AEST	657	657	1	A/N	NOTE 3.
	1M3A	TOTAL MALE AIDES	658	663	6	N	NOTES 1 & 2.
		1M3BEST	664	664	1	A/N	NOTE 3.
	1M3B	TOTAL FEMALE AIDES	665	670	6	N	NOTES 1 & 2.
		1M3CEST	671	671	1	A/N	NOTE 3.
	1M3C	TOTAL AIDES	672	677	6	N	NOTES 1 & 2.
		1NAEST	678	678	1	A/N	NOTE 3.
	1NA	MALE OFFICE/ CLERICAL PERSONNEL	679	684	6	N	NOTES 1 & 2.
		1NBEST	685	685	1	A/N	NOTE 3.
	1NB	FEMALE OFFICE CLERICAL PERSONNEL	686	691	6	N	NOTES 1 & 2.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page 25
Subsection(s) .01
Date Documented 10/26/78
[X] New [ ] Revision
Change Notice #

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX KEYTAPING

212.01 KEYTAPING INSTRUCTIONS

CHARGE # SYSTEM ID: CCD PAGE 12 OF 17
JOB TITLE DATE: 10/26/78
SOURCE DOCUMENT PART VI LEA NON FISCAL REPORT FORM # 2393-2 (5-1-78)
COMMENTS

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		1NCEST	692	692	1	A/N	NOTE 3.
	1NC	TOTAL OFFICE/ CLERICAL PERSONNEL	693	698	6	N	NOTES 1 & 2.
		10AEST	699	699	1	A/N	NOTE 3.
	10A	OTHER MALE NON-PROFESSIONAL PERSONNEL	700	705	6	N	NOTES 1 & 2.
		10BEST	706	706	1	A/N	NOTE 3.
	10B	OTHER FEMALE NON-PROFESSIONAL PERSONNEL	707	712	6	N	NOTES 1 & 2.
		10CEST	713	713	1	A/N	NOTE 3.
	10C	TOTAL OTHER NON-PROFESSIONAL PERSONNEL	714	719	6	N	NOTES 1 & 2.
		1PAEST	720	720	1	A/N	NOTE 3.
	1PA	TOTAL OF ALL MALE PERSONNEL	721	726	6	N	NOTES 1 & 2.
		1PBEST	727	727	1	A/N	NOTE 3.
	1PB	TOTAL OF ALL FEMALE PERSONNEL	728	733	6	N	NOTES 1 & 2.
		1PCEST	734	734	1	A/N	NOTE 3.
	1PC	TOTAL OF ALL PERSONNEL	735	740	6	N	NOTES 1 & 2.
		2AEST	741	741	1	A/N	NOTE 3.
	2A	NO. PREKINDERGARTEN PUPILS	742	747	6	N	NOTE 2.

OFFICE OF EDUCATION

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page [illegible]

Subsection(s) .01

Date Documented 10/26/78

[X] New [ ] Revision

Change Notice # _____

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX KEY [illegible]

212.01 KEY [illegible] INSTRUCTIONS

CHARGE # _____ SYSTEM ID: CCD PAGE 13 OF 17

JOB TITLE _____ DATE: 10/26/78

SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT FORM # 2393-2 (5-1-78)

COMMENTS _____

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		2BEST	748	748	1	A/N	NOTE 3.
	2B	NO. KINDERGARTEN PUPILS	749	754	6	N	NOTE 2.
		2CEST	755	755	1	A/N	NOTE 3.
	2C	NO. FIRST GRADE PUPILS	756	761	6	N	NOTE 2.
		2DEST	762	762	1	A/N	NOTE 3.
	2D	NO. SECOND GRADE PUPILS	763	768	6	N	NOTE 2.
		2EEST	769	769	1	A/N	NOTE 3.
	2E	NO. THIRD GRADE PUPILS	770	775	6	N	NOTE 2.
		2FEST	776	776	1	A/N	NOTE 3.
	2F	NO. FOURTH GRADE PUPILS	777	782	6	N	NOTE 2.
		2GEST	783	783	1	A/N	NOTE 3.
	2G	NO. FIFTH GRADE PUPILS	784	789	6	N	NOTE 2.
		2HEST	790	790	1	A/N	NOTE 3.
	2H	NO. SIXTH GRADE PUPILS	791	796	6	N	NOTE 2.
		21AEST	797	797	1	A/N	NOTE 3.
	21A	NO. ELEMENTARY 7TH GRADE PUPILS	798	803	6	N	NOTE 2.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page 212-01 17
Subsection(s) .01
Date Documented 10/26/78
X New ___ Revision
Change Notice # ___

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX KEYTAPE

212.01 KEYING INSTRUCTIONS

CHARGE # ___ SYSTEM ID: CCD PAGE 14 OF 17
JOB TITLE ___ DATE: 10/26/78
SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT FORM # 2393-2 (5-1-73)
COMMENTS ___

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		2JAEST	804	804	1	A/N	NOTE 3.
	2JA	NO. ELEMENTARY 8TH GRADE PUPILS	805	810	6	N	NOTE 2.
		2OAEST	811	811	1	A/N	NOTE 3.
	2OA	NO. ELEMETARY UN-CLASSIFIED PUPILS	812	817	6	N	NOTE 2.
		2PAEST	818	818	1	A/N	NOTE 3.
	2PA	TOTAL ELEMENTARY PUPILS.	819	825	?	N	NOTE 2.
		2IBEST	826	826	1	A/N	NOTE 3.
	2IB	NO. SECONDARY 7TH GRADE PUPILS	827	832	6	N	NOTE 2.
		2JBEST	833	833	1	A/N	NOTE 3.
	2JB	NO. SECONDARY 8TH GRADE PUPILS.	834	839	6	N	NOTE 2.
		2KEST	840	840	1	A/N	NOTE 3.
	2K	NO. 9TH GRADE PUPILS. SECONDARY	841	846	6	N	NOTE 2.
		2LEST	847	847	1	A/N	NOTE 3.
	2L	NO. 10TH GRADE PUPILS SECONDARY	848	853	6	N	NOTE 2.
		2MEST	854	854	1	A/N	NOTE 3.
	2M	NO. ELEVENTH GRADE PUPILS. SECONDARY	855	860	6	N	NOTE 2.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page 212-01 28

Subsection(s) .01

Date Documented 10/26/78

X New ___ Revision

Change Notice # ___

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX KEY PUNCH INSTRUCTIONS

212.01 KEY PUNCH INSTRUCTIONS

CHARGE # ___ SYSTEM ID: CCD PAGE 15 OF 17

JOB TITLE ___ DATE 10/26/78

SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT FORM # 2393-2 (5-1-78)

COMMENTS ___

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		2NEST	861	861	1	A/N	NOTE 3.
	2N	NO. 12TH GRADE PUPILS. SECONDARY	862	867	6	N	NOTE 2.
		2OBEST	868	868	1	A/N	NOTE 3.
	2OB	NO. UNCLASSIFIED SECONDARY PUPILS.	869	874	6	N	NOTE 2
		2PBEST	875	875	1	A/N	NOTE 3.
	2PB	TOTAL SECONDARY PUPILS.	876	882	7	N	NOTE 2.
		3AEST	883	883	1	A/N	NOTE 3.
	3A	NO. MALE 12TH GRADE GRADUATES.	884	888	5	N	NOTE 2.
		3BEST	889	889	1	A/N	NOTE 3.
	3B	NO. FEMALE 12TH GRADE GRADUATES.	890	894	5	N	NOTE 2.
		3CEST	895	895	1	A/N	NOTE 3.
	3C	TOTAL NO. 12TH GRADE GRADUATES.	896	900	5	N	NOTE 2
		4AEST	901	901	1	A/N	NOTE 3.
	4A	NO. PUBLIC SCHOOL PUPILS TO BE BUSSED.	902	906	5	N	NOTE 2.
		4BEST	907	907	1	A/N	NOTE 3.
	4B	NO. PRIVATE SCHOOL PUPILS TO BE BUSSED.	908	912	5	N	NOTE 2.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page 212-01 29.
Subsection(s) .01
Date Documented 10/26/78
[X] New [ ] Revision
Change Notice #

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX KEY [illegible]

212.01 KEY [illegible] INSTRUCTIONS

CHARGE #_____ SYSTEM ID: CCD PAGE 16 OF 17

JOB TITLE_____ DATE: 10/26/78

SOURCE DOCUMENT PART VI LEA NONFISCAL REPORT FORM # 2393-2 (5-1-78)

COMMENTS_____

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		4CEST	913	913	1	A/N	NOTE 3.
	4C	TOTAL NO. PUPILS TO BE BUSSED.	914	918	5	N	NOTE 2.
		5AEST	919	919	1	A/N	NOTE 3.
	5A	NO. LARGE TRANSPORTATION VEHICLES OWNED.	920	923	4	N	NOTE 2.
		5BEST	924	924	1	A/N	NOTE 3.
	5B	NO. SMALL TRANSPORTATION VEHICLES OWNED.	925	928	4	N	NOTE 2.
		5CEST	929	929	1	A/N	NOTE 3.
	5C	TOTAL TRANSPORTATION VEHICLES OWNED.	930	933	4	N	NOTE 2.
		6EST	934	934	1	A/N	NOTE 3.
	6	SQUARE MILES WITHIN AGENCIES BOUNDARY.	935	941	7	N	NOTE 2.
		7EST	942	942	1	A/N	NOTE 3.
	7	NO. OF MEMBERS ON THE BOARD OF EDUCATION	943	945	3	N	NOTE 2.
		8EST	946	946	1	A/N	NOTE 3.
	8	NO. OF SCHOOLS OPERATED BY THIS AGENCY ON OCT.	947	949	3	N	NOTE 2.
		9EST	950	950	1	A/N	NOTE 3.
	9	NO. OF SCHEDULED DAYS PUPILS EXPECTED TO ATTEND	951	953	3	N	NOTE 2.

OFFICE OF EDUCATION

DATA PROCESSING SYSTEM DOCUMENTATION AND OPERATING PROCEDURES MANUAL

Section & Page: 10

Subsection(s): .01

Date Documented: 10/26/78

[X] New [ ] Revision

Change Notice #:

SYSTEM TITLE:	SYSTEM ID:
COMMON CORE DATA	CCD

212.XX KEYTAPE

212.01 KEYLINE INSTRUCTIONS

CHARGE #: SYSTEM ID: CCD PAGE 17 OF 17

JOB TITLE: DATE: 10/26/78

SOURCE DOCUMENT: PART VI LEA NONFISCAL REPORT FORM #: 2393-2 (5-1-78)

COMMENTS:

CARD NUM	REF. NO.	FIELD TITLE	COLUMNS FR	COLUMNS THR	COLUMNS NO.	A/N	INSTRUCTIONS
		NEW CLOSD	954	954	1	A/N	BLANK = SCHOOL IS OPEN N=NEW SCHOOL. C=CLOSED.
		BLANK	955	1000	46	A/N	LEAVE BLANK.
NOTE # 1. FIELDS AAA THROUGH APC WILL BE ROUNDED TO THE NEAREST TENTH. THE DECIMAL WILL BE UNDERSTOOD BUT NOT PRESENT ON TAPE.							
NOTE # 2. NUMERIC FIELDS WILL BE RIGHT JUSTIFIED AND ZERO FILLED TO THE LEFT. ZERO FILL THE FIELD IF DATA WAS NOT PROVIDED.							
NOTE # 3. THE ESTIMATION INDICATOR FIELDS CONTAIN THE SUFFIX "EST" IN THEIR NAMES. DATA FIELDS THAT ARE ESTIMATED WILL BE ANNOTATED BY AN "*" ON THE FORM. THE "*" SHOULD BE KEYED IN THE ONE DIGIT FIELD AT THE LEFT OF THE APPROPRIATE DATA FIELD ON TAPE. WHEN N/A APPEARS ON THE FORM PLACE AN "N" INTO THE APPROPRIATE ESTIMATION INDICATOR AND ZEROS INTO THE DATA FIELD. FOR TRUE DATA PLACE A ZERO OR A SPACE INTO THE APPROPRIATE ESTIMATION INDICATOR.							